PROCEEDINGS FROM THE ETS & ARI EMOTIONAL INTELLIGENCE WORKSHOP

Richard D. Roberts & Jennifer Minsky (Eds.)

VENUE

Chauncey Conference Center, Princeton, New Jersey, U.S.A.

DATE

13-15 NOVEMBER, 2003

ORGANIZERS

Richard D. Roberts, ETS (Chair)
Paul A. Gade, ARI
Patrick C. Kyllonen, ETS
Moshe Zeidner, University of Haifa, Israel
Gerald Mathews, University of Cincinnati
William Strickland, Human Resources Research Organization

Volume II

Sessions III and IV

Report Documentation Page

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 MAR 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Proceedings from the ETS and Army Research Institute Emotional Intelligence Workshop Volume 2, Part 2: Related Constructs	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-BR 2511 Jefferson Davis Highway Arlington, VA 22202-3926	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001936., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**404**	

VOLUME II

TABLE OF CONTENTS

SESSION III: RELATED CONSTRUCTS
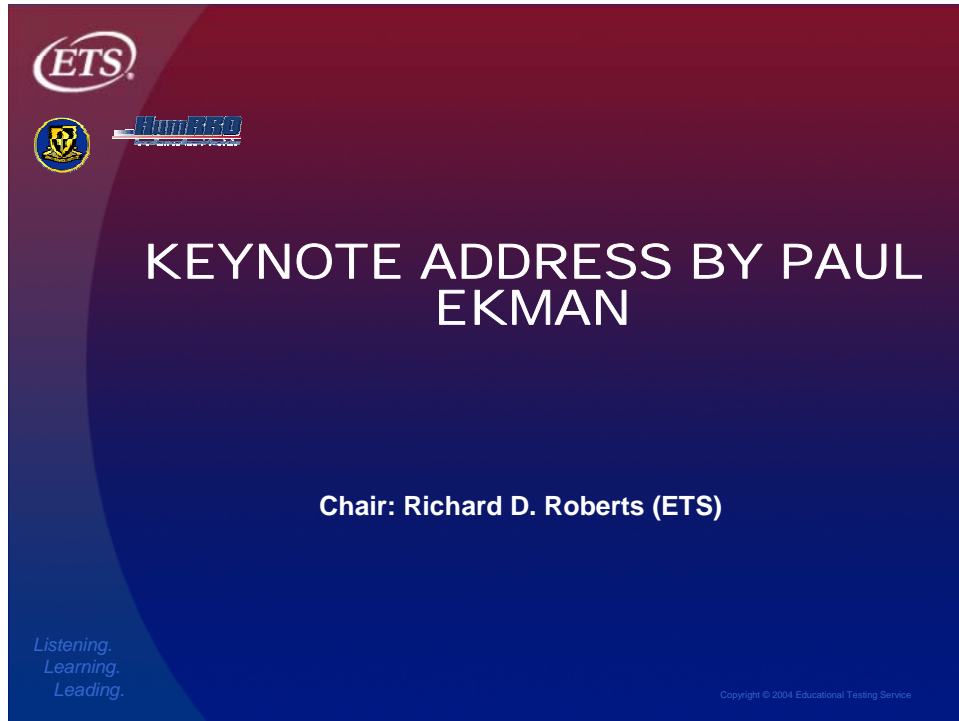
DR. ROBERTS: I really do want to welcome today, Paul Ekman, he is actually a legend! That's it.

Emotional Skills

Paul Ekman
University of California at
San Francisco

EMOTIONAL SKILLS

DR. EKMAN: I want to thank you for this opportunity to give this talk today. It's mostly new ideas and some new findings. I also want to apologize to some of the other speakers, some of whose talks I've missed. I'm still recuperating from major surgery and I have to take breaks each day, so it's unavoidable that I can't be at all the talks, much to my regret.

I want to begin by describing what the consensus is within psychology about what an emotion is and then I will discuss each of four skills. Klaus Scherer yesterday objected to the word "intelligence", I will object to the word "emotional" because I'm not at all sure that the people concerned with emotional intelligence are actually focusing on emotion. I think most of them are not focusing on what I'll describe as the general consensus among emotion theorists and researchers about what emotion is. Nevertheless, there are skills relevant to emotion that might be relevant to emotional intelligence. That is my reach for the topic today.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Almost 10 years ago Richard Davidson and I co-edited a book called, "The Nature of Emotion" in which we invited 24 emotion researchers and theorists to write their views on issues that we knew there was disagreement about. These are half of all the 12 questions we asked them to answer about emotion and as you can see, they represent not orthodoxy, but quite a different range of views. For each question we picked five or six people, who we knew pointedly disagreed, to write specifically on that question. Here are the rest of the questions.

Why am I telling you this? Because in the epilogue to the book, Davidson and I while looking across all of this, identified what these people agreed about. First, it's that emotions involve information processing. You heard Klaus Scherer talk about his approach to appraisal; we get along so well because he looks at what precedes emotion and I look at emotional responses. But today, I'm going to invade his area at least in the first part of my talk. Information processing involves an evaluation typically of some event that precedes the emotion, and that it involves expressive and physiological changes distinctive for each emotion. But, there was disagreement about both the number of emotions and whether it's a universal phenomenon.

Emotional Skills

Paul Ekman
University of California at
San Francisco

However, there was no disagreement about the memories, expectations, methods of coping, or that there is subjective experience. In terms of the signal, which is what I've spent most of my life studying, there is disagreement about how many emotions have a distinctive signal and secondly, whether the signal is the source. That is, can there be emotions that have no signals, are those still considered emotions? In addition, there was clear consensus concerning mood, namely that moods last a lot longer and they often, but not always, have no obvious social cause.

Incidentally, when considering the history of the study of emotion, there is really a gap in the English speaking world between 1924, Floyd Alport's chapter in emotion, and Solomon Ash's chapter in the middle 50's. Emotion research pretty much died; behaviorism killed it in America. Expression psychology in Europe and particularly in Germany was quite active until Hitler killed it when he had the social scientists who didn't flee, give it a racist bias, so that you could identify inferior races from their expressions and gestures. So, there was a big gap. If you look at introductory textbooks in the 1950s, not only is emotion not a chapter, it's not even in the index. It just didn't exist; people didn't get taught about it. "If I can't measure it, it doesn't exist", has been sort of the motto of psychology. I think that motto is still around, but I think the tide changed in part because we can now measure at least the rear end if not the front end.

Emotional Skills

Paul Ekman
University of California at
San Francisco

What about the disagreement about how many emotions are there? There seems to be general consensus that there are about seven, however, there is disagreement about love, hate, jealously and compassion. None of those do I consider an emotion. I would be glad to tell you why in the question period; I'm not going to spend time on it. That doesn't mean they don't exist, but not everything that is important or has feelings is an emotion.

Is there a goal standard? Well, some have thought that subjective experience is the goal standard - if you are not aware of your emotion you don't feel it. No one that has ever done clinical work believes that, I don't think anyone who has been married believes that. Your spouse doesn't always know they are feeling emotional and they act in a way that by all outside objective views is very emotional. Is physiology the gold standard? Well, I don't know a physiological psychologist who doesn't think it is the goal standard. Is expression the goal standard? Well, at one point I took that view, I no longer do, because I don't think there is a gold standard. As Klaus said yesterday, there are many different characteristics that combine to distinguish emotion from related affected states; emotion is only one affective state.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Now let me turn to emotional skills. Most of what I'm going to talk about is based on a new book called "Emotions Revealed", that was published last spring and is just about to come out in French, Spanish and Chinese. First, I should say that I don't claim that the four emotional skills are the only ones, they are just four that I have considered. I think they are four important ones, but clearly there are other emotional skills. The first skill is to be aware of the emotion we are feeling, ideally as it is developing but at least in the first few moments before much or any action is taken or any words are spoken. This is not easy to achieve. I believe that this is so, because I don't think we evolved in order for consciousness to muck around with our emotions. Again, as Klaus mentioned yesterday, our emotions save our lives because we don't need to think about things first in order to become emotional. If, for example, you think of a near miss car accident, that very complex process of evaluating that car that is about to hit you and determining how you are going to manifest the behavior to cope with it, that all occurs prior to consciousness. You become conscious once you're already doing it often after it's already over, if you have survived. And if you had to think about it, you probably wouldn't survive. And that's when emotions really are in our service. Now, my conscious mind was involved in learning how to walk, but I didn't think of one step I took today. When I learned how to eat, I was very conscious of each movement of that fork, but I didn't think a bit about the movements of the eating implements at breakfast this morning. The fact that behavior is totally automatic without thought does not necessarily mean that it's biologically built into us, but it does mean that it occurs without conscious deliberation or awareness.

Emotional Skills

Paul Ekman

University of California at

San Francisco

Now, most emotion researchers agree that the appraisal process that turns on emotion often typically operates quickly outside of awareness and too fast for awareness to penetrate as it's occurring. There are eight other ways, I'm not going to focus on them, but let me just put them on the board. These are different ways in which emotions can be brought about. Apart from the automatic appraisal which I think is the one that most often, both saves our life and gets us into trouble, and is the most common one.

Let me turn away from that and come back to this first skill. Once we become emotional we are often gripped by it. We behave without consideration of how we should enact our emotion. It's not that we think about, "Well should I challenge you on the insult you just gave me or should I just let it pass over?" That can happen, but it often it doesn't happen. The first skill is to know that an emotion is occurring as early in the episode as possible. Sometimes our automatic appraisal is on target- it is a snake not a coiled rope, but sometimes it isn't. Obviously, those who responded to things that looked snake-like probably put more genes into the gene pool than those who thought they were just passing a coiled rope and got bit. Sometimes we do misunderstand what's occurring in our emotions. They are either too strong or they are inappropriate or they're too weak. And this happens to anyone. And the antidote is to be aware that you are becoming or being emotional as early in the process as possible.

Emotional Skills

Paul Ekman

University of California at

San Francisco

Now even the Dalai Lama, who I've had the good fortune to be able to meet with three times, does not claim that he can be aware of the automatic appraisal that initiates an emotional episode. But through hard work by developing skills that we don't start out with, some people can learn impulse awareness. That's the first skill, it's impulse awareness - that means being aware of the impulse before the action, the impulse before the words are spoken. It's very tough; I'm not convinced that everybody can learn that skill. Shortly, I'll describe the kinds of people who are most equipped to be able to learn it and some of things you can do to enhance that skill. Now, just so you shouldn't think that this is all coming out of Tibet and Buddhism, 45 years ago when I was still in the grip of clinical psychology, my psychotherapy supervisor said, "If you can increase the gap between impulse and action, you will have enormously benefited your patience." That is a very common view in psychodynamics, psychotherapy. The Buddhist put it, "To recognize the spark before the flame"; it's exactly the same concept. Now notice it's not recognizing what gives birth to the impulse; what starts the spark. That, virtually no one thinks you can enter into. But to be aware and not necessarily let the flame occur, that's the first extremely difficult, but extremely useful skill.

Emotional Skills

Paul Ekman
University of California at
San Francisco

There are a lot of things you can do to try to enhance this, let me just mention a few of them. One is to keep a diary of regrettable, and just so you won't feel too badly, non-regrettable emotional episodes. And then try to identify from the list of regrettable emotional episodes what's the common appraisal that leads you to these regrettable episodes. You can usually, as Klaus and researchers have shown, figure out afterwards what it was that you've misappraised. What script did you bring in that doesn't fit the situation, if you want to use that terminology. Typically, for laymen, it's easier to think of it as a trigger than appraisal but the appraisal is making something into a trigger. In psychotherapeutic terms, one can work through the script and lay it to rest through the interaction of the therapist. My bet is that that never succeeds completely. The most one can do is to learn to recognize and weaken that particular appraisal process; much of emotional behavior and appraisal is learned. Once learned, is not unlearnable. I think of emotion as a fish trap, it's easy to get in but hard to get out. People spend a lot of money trying to get things out of the system that have become automatic, I think you can weaken them but not eliminate them.

Emotional Skills

Paul Ekman

University of California at

San Francisco

A totally different approach to achieving impulse awareness that I'll describe later, is mindfulness meditation. The Buddhists have developed exercises in which you focus your attention exclusively on each breath you take and the sensation in your nostrils as the breath goes in and out. Or, on each step each you take. If you've ever seen people do walking meditation, they look like zombies because they are walking slowly. Or, eating meditation is really good for people who want to lose weight, because it takes you about five minutes to get each mouthful. There isn't any hard scientific evidence at this point that mindfulness meditation improves emotional life, although there are many, many studies in which people who've done it claim in self-report measures that their emotional life has been improved. Up until very recently I couldn't understand why focusing on breathing would benefit emotional life, yet I had seen so many people who acted in ways in which I could not act who had been meditators for 10 or 12 years.

Emotional Skills

Paul Ekman
University of California at
San Francisco

One of the things they shared in common was seeming to be aware of the emotions that were occurring much earlier in the interaction than I ever was. About a month before I wrote the notes for this presentation, like a bolt out of the blue, it struck me why mindfulness meditation might work. It is the very practice of learning to focus on an automatic process that requires no conscious monitoring that creates a capacity to be attentive, I think, to other automatic processes. We breathe without thinking; I presume none of you have thought about a breath you've taken since you've woken up this morning. You don't need to, nature doesn't require it. Think what life would be like if you have to think about each breath or you would die. And so, when you try to think about each breath without letting your consciousness drift to anything else, it's extremely difficult. The first time I tried this, I could not manage. One of the techniques I used was to simply count and see how far I could get into the count before my mind drifted on to something else. And in a full day, I never got to sixty. The Buddhist teacher said, "That will take weeks and months to be able to have a full minute where all you're doing is focusing on this automatic process, nature didn't provide us the means to do that." But it is possible to learn by a lot of practice and my suggestion is that when you learn that skill, or any skill, you are creating new neural pathways. You are creating neural pathways that nature didn't give you, neural pathways for monitoring automatic behavior and they have the ability to transfer. So, if you can monitor your breathing you are going to be more aware of the automatic processes when that impulse is generated.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Talking to my buddy, Richie Davidson, who works on the brain and emotion, he thinks it makes sense and talking to the Dalai Lama a month ago, he thinks it makes sense; the Buddhists hadn't considered before why it is helpful to emotions. The fact that people think something makes sense doesn't mean it's true, but it is at least suggestive. A research project is now going on at UCSF that I organized and is being led by Margaret Kennedy, in which very hard objective measures of emotion, not only self-report is being used, to look at before and after on the impact of mindfulness meditation.

Here is the second skill. Not as good as the first, but the impulse has been transformed into actions and words. You are exercising, maybe not in the first instant but before it has gotten very far, how you are in that emotion. Again, emotions did not evolve, in my view, for you to be able to do this. It is not going to be easy for you to do this; it really requires learning a skill and learning a skill that is not like learning to ride a bicycle, bicycles were not part of our ancestral environment. And yet it is easy to learn a bicycle, both of my kids learned it in less than an hour. I hadn't ridden a bicycle for 25 years, and I got right back on and rode a bicycle. Once you learn this novel piece of behavior, it sticks with you. The skills I'm talking about do not. They don't connect to anything within our brain that allows them to be maintained, unlike riding a bicycle which connects to all kinds of things having to do with locomotion and visual motor coordination. What you want to be able to do is pause and consider "Is this the right emotion for this situation?" And am I acting it in a way that is going to be most useful to my goals and to my relationship with the other person?

Emotional Skills

Paul Ekman

University of California at

San Francisco

Now, it's not that you are unconscious when you're emotional, you are just not conscious of being emotional. A memory researcher, I can't remember his name, gives the example, which I think is an excellent one, of you're reading a book and you're turning the pages and after a few minutes you notice the fact that you've been thinking about the movie you saw last night and you didn't read a word on any of the last five pages. When you were thinking about that movie you were not unconscious. You knew you were thinking about a movie- you were conscious, but you weren't conscious of the fact that that's what you were doing, that's why you continued to turn the pages. That has a lot of applicability to emotional behavior. When we're emotional, it's not that we aren't conscious, we're very conscious, but we're not conscious to the fact that we are being emotional. The Buddhists call that a "watcher", there's nothing watching you, there's no meta-consciousness; there's no attentiveness to the fact that you're being emotional.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Now everyone reacts quickly in a near miss car accident but some of us react to all of life like if it is a near miss car accident. One of the components of what I call our emotional profile is how quickly we respond emotionally. Some of us have a very fast emotional response system, while for others the emotional response system seems to be slower. And in examining the latency of facial expression to emotionally arousing films, I have found that the speed of response is a general, rather than an emotion specific, phenomenon. Our language distinguishes such differences in latency for anger - a short fuse versus smoldering. But what we found in research, and I'll shortly summarize, is if you have a short fuse for anger you have a short fuse for fear, sadness, and for disgust. It doesn't matter how the emotion is elicited in terms of the circumstance or what the emotion is; most people show consistency in emotional latency. There is another reason why this skill is very hard to acquire and I'm proposing that when we become emotional there is initially a refractory period. And during that refractory period, information that contradicts the emotional feeling is not accessible to us, either from the environment or stored in our own knowledge banks. All that we filter is things that support the emotion; it's a focusing. It's great if it lasts a couple of seconds but a disaster if it lasts for minutes, because nothing that can disconfirm it can get into your consciousness.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Now what can we do to obtain emotional behavior awareness, skill number two? In my book, I describe exercises that we are now testing in the project I mentioned earlier, exercises in which you self-induce emotion using one of three techniques that have been our favorites in laboratory studies. Then, you focus your consciousness on the physiological changes in your body. So that those become a louder signal that tells you, "I'm starting to get emotional." They're there and some of them are capable of being reported to consciousness, but for most of us it is pretty dim, so the idea is to heighten it. The second technique is the identification of a regrettable emotional episode. And the third technique is heightening your sensitivity to the emotional reactions of others. In other words, I don't get to see my face, but you do, so you have more information in some sense, particularly if I'm not speaking. So, you get information about what is going on from me from my face, and I don't; there is very little facial feedback. We have experimental evidence to support the idea that most people can't tell what their face is doing. You get great feedback for temperature and touch, but not for muscular movement. But from the other person's reactions, they can inform you of what's going on in you, if you become more sensitive. So, it's not simply that they'll enable you to better deal with the other person; they enable you to acquire this skill, to realize you are becoming emotional. Not everyone will benefit equally from these exercises. I mentioned one reason a few moments ago, which is whether your emotional profile responds slowly or quickly. Those who respond more slowly have a better chance of being able to acquire and use this skill. There is a second feature of emotional profile which enters in, which is the magnitude of your emotional response.

Emotional Skills

Paul Ekman
University of California at
San Francisco

In collaborative research with Bob Levinson, for the last 10 years we have been slowly analyzing a data set that involves three response systems - the nervous system activity, facial muscular movements and subjective experience; three emotions- fear, sadness, and disgust; and two different emotional eliciting circumstances- viewing short film clips and reliving a past emotional experience. The reason we're limiting it to these three is because we have really not been able to find films for other emotions that are emotion specific. It is essentially a 2X3X3 design, but don't expect to see the ANOVA; the data analysis will be complete this winter. So we have two circumstances, films and memory, with three emotions with three kinds of emotional responses. And we found generality across responses, across emotions, and across eliciting circumstances in this strength of emotional response. And I know unpleasant things were said about factor analysis but when you do a factor analysis we found that when you put the facial measures, the autonomic and the subjective in, they all load on the same factors. Another way of putting that more simply is if you just look at straight means or you look at intercorrelations. You can take an aggregate measure of autonomic activity that correlations with an aggregate measure of the strength of facial muscular contractions which correlates pretty highly with ratings on the good scales. So, the strength of our emotional response is something that seems to be for most people.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Another thing that is really quite interesting is that people are quite aware of this characteristic and a simple questionnaire that asks people to rate the strength of a typical emotional experience in their daily life, given to them three weeks before they came into our laboratory, predicts better than .4 the strength of facial autonomic responses. It doesn't predict the self-report very well, but it does report physiological and expressive behavior across these three emotions and across these two tests. So if you are a big responder, you know it, people are letting you know. And if you are a very small responder then probably your spouse is saying, "Why don't you seen to care about things the same way I do." Another really amazing thing is that the people who study marital interactions have found that people who have different emotional profiles often marry each other. They then spend a lot of time trying to convince the other person, you should have the profile I have, it's the right one and yours is the wrong one.

Emotional Skills

Paul Ekman
University of California at
San Francisco

So if we look at this, the strength of response and the speed of response, it turns out that all four cells are filled. That is, there are people who are slow, moderate responders; I said that they should make good negotiators or mediators, which is the hypothesis. What is not a hypothesis is that those people exist. Furthermore, there are people that are fast, extreme responders; these are the attack dogs of life. There are uses for attack dogs but attack dogs don't work well as negotiators or mediators. Is this modifiable? Nobody knows. My bet is that this is not modifiable except through trauma. And what trauma does is move everybody into cell four. Now, it could be that if we were to design traumatic experiences of ecstasy, that is traumatic in that they are enduring, they last for many, many hours and they are spectacularly overwhelmingly enjoyable, if we could design such environments, maybe we could make other shifts. But one way of thinking of PTSD, is it moves people into cell four.

Emotional Skills

Paul Ekman
University of California at
San Francisco

The third emotional skill is becoming more sensitive to how others feel. I mentioned two good reasons for doing that. One, is so you better understand the other person and can better calibrate your own responses. The second, is so you better understand what is going on in you. I am going to emphasize the face, but I do believe that voice is every bit as important as the face; it's different in interesting ways. I just stopped the voice, I pause it, no signal. I can't stop the face, it's always there. The voice reaches people who aren't looking at you. Can you imagine what it would be like to do child care if there weren't vocal signals? You would always have to be looking at your kid, it would be impossible, you would be totally preoccupied. But hunters can't communicate with their voice because they would give away their location to the prey. So, both of these are overlapping and highly redundant. It's only during very explicit deception situations that we have found disconjunctions; they are highly overlapping and consistent. But we are now in the position that we can really teach people about facial recognition sensitivity.

Emotional Skills

Paul Ekman
University of California at
San Francisco

So let me show you, what we're doing. Seven emotions, grouped into five and two, because the five there has the strongest evidence and for the two, there is still some argument about. I don't think there is much argument, but I have to acknowledge that some others do and we are just resubmitting our fourth rebuttal of another article that says that contempt doesn't have a unique signal. Actually, the evidence in Indonesia is the best in the world, what an unexpected finding, but I think just because of the linguistic representation. When you use a story, which face fits with a person who claims to have accomplished things that other people did not, then you get extremely high agreement across cultures in identifying what we said is the contempt face. But the word itself, particularly in English, can you think of a worse word than contempt? Our research and that of many others have shown that people worldwide have no problem and need no aid in interpreting a facial expression of emotion, if it is seen in isolation and it is one of these seven. If it is seen with accompanying speech and it doesn't contradict this speech content, then no one needs tutoring. But life isn't often like that.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Incidentally, Darwin proposed on the basis of one data point, that the ability to recognize emotion was innate, not just the expression. His one data point was his son William, his first of 10 children. And Darwin's two keys to success were he never had a job in his life and he never gave an academic lecture. He observed and kept a daily diary on each and every one of his 10 children and he did experiments with them. When William was between two and three, and William had never seen another child at that point, Darwin asked the nanny to put a look of distress on her face and to show a few sobbing movements. William immediately went and put his arm around her to comfort her. So, Darwin said he'd never seen this behavior before, yet he knows what it is. He knows what it is, not just because he was able to give us a rating on pleasantness and unpleasantness or choose sadness from one of the lists, but because he showed the appropriate behavior toward someone who was sad. When expression contradicts speech content then our evidence, based only on studies in English speaking countries, is that people don't use the information in the face, instead they emphasize speech content. Even if you tell people that the people they are about to see may be lying to them, they still emphasize mostly speech content and ignore the valuable information in the face and the voice that could be informative to them when particularly looking at discrepancies.

Emotional Skills

Paul Ekman
University of California at
San Francisco

So, we've developed two self-instructional CDs to improve emotional skills. One is to teach people how to recognize micro-facial expressions. In 1969 we published our first article about micro-facial expressions and so did Haggard and Isaacs. We said that they are the consequence of deliberate concealment and Haggard and Isaacs said they are the consequence only of repression. I know we're right because we have a lot of experimental data and at least from looking at clinical interviews from psychoanalysts, I have been able to confirm at least three or four examples. There is every reason to think that the person is totally unconscious and you see the micro expression. You can't tell from the micro expression then whether the concealment is conscious or unconscious, but you can tell that an emotion is being concealed.

I am going to show you an example of one so you won't think this is something I made up in the laboratory, and I've chosen Kato Kalen. The Europeans here may not know who this famous person is, he was sort of a sycophant for O.J. Simpson, he testified in the trial. Kato had to testify because he was not on trial and Marsh Clark, the prosecutor, was badgering him and at this point had caught him in a lie. He had said he had not negotiated a book advance but she knew he had from the book publisher. And this later came out, and she's pushing him because she's trying to destroy his total credibility to the jury.

Emotional Skills

Paul Ekman
University of California at
San Francisco

MR. CLARK: Mr. Kalen you got a lot of money for your appearance on Current Affair didn't you?"

MR. KALEN: Um, yes.

MR. CLARK: And as a matter of fact, isn't it true Mr. Kalen that you have a book proposal out for about half a million dollars right now don't you?

MR. KALEN: No.

MR. CLARK: You don't have a book proposal?

MR. KALEN: No.

MR. MARSH: Aren't you represented by the William Morris Agency to write a book?

MR. KALEN: No.

MR. MARSH: Haven't you written a book proposal that has been submitted to St. Martins for publication?

MR. KALEN: No.

MR. MARSH: You are not aware of any contract for half a million dollars that you have signed.

MR. KALEN: I know it's out there but I haven't done that.

MR. MARSH: You haven't signed it yet?

Emotional Skills

Paul Ekman
University of California at
San Francisco

MR. KALEN: Oh no, I heard a story about that, that's not true.

MR. MARSH: You don't have any book proposals out?

MR. KALEN: No. Don't want to do a book.

MR. MARSH: Do you plan to write a book in the future Mr. Kalen?

MR. KALEN: As of today, no way.

DR. EKMAN: Incidentally, just to show you that when we teach people about detecting deception from demeanor, we don't emphasize just the face, of course. He showed what we just call soft voice and it takes some skill on the part of the interviewer to get them into a position where they either have to lie or tell the truth. Some people, not everyone, do a soft voice, just as he did.

Okay, so now I'm going to show you the same thing again but first you'll get a normal speed then slow-motion version of the micro expression that was contained in that segment. Notice that like all of the micro expressions that we've identified it is very fast; this one is about 1/20th of a second, it is very extreme. If it was on the face this long, no one would have trouble but the problem is that it isn't.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Now the question is can you teach people to recognize it? We know from our research that nearly everybody misses them. Danny Goleman put in his book, "Destructive Emotions" our finding three years ago, which we haven't published, that two Buddhist monks did spectacularly well in identifying micro expressions and the Buddhist world is very excited about that. They were a little discouraged to hear that we can now teach anybody in 25 minutes to do just as well.

This is the micro expression training tool that we developed and I'll show you how it works and then some results from it. Now this tool has a number of parts. It has a pretest in which you see 14 micro expressions and it computes your accuracy score, so you know where you are to begin with. Then, there is a training segment, which I'm going to show you an example from where emotions that are often confused with each other are shown in a morphed version with commentary. Then, there is a practice where you get to give yourself feedback, make judgments and find out whether you're right or wrong. And then there is a review that again, contrasts emotions which are easily confused. And finally, there is a posttest. So, I'm skipping the pretest because of time. Now, here is the first of the training segments:

Emotional Skills

Paul Ekman
University of California at
San Francisco

"Anger and disgust are sometimes confused with each other but you can see they are very different. Although the brow comes down in both, watch what happens in the eyes. On the left, in anger, the eyes are glaring, they are narrowed. And on the right, not glaring. Also, watch in the next time it cycles around, the lips. See how tightly pressed they are? But they are relaxed in disgust. All of the action in disgust is in the center line."

Now, you could say, why am I using such extreme expressions? Because that's what you see in micro expressions, they are typically just as strong as this, as you saw in Kato Kalen. So this is fear and surprise, this is sadness and fear, this is happiness and content. And after you have gone through that, and you can see this is all set up so it's self instructional, you don't need a teacher. We used to teach this, now we don't, we just give people the CD.

The practice session works like this, let me just show you an item or two from it. There are 28 faces, and incidentally throughout this you never see the same person twice. You now you have to say, what emotion was that? So let's suppose you say you thought she was angry, and you were wrong, you get feed back. You can try disgust; it tells you that you were right. So, you get to do that 28 times then you get a review, this whole training takes about 24 minutes. We get about a 40% increase from pre to post accuracy.

Emotional Skills

Paul Ekman
University of California at
San Francisco

In collaboration with Mark Frank at Rutgers we collected data on Coast Guard criminal investigators, and randomly assigned them to a training or a control group. The training group got the training that we just covered, the control group did not. And then we showed them eight videotaped clips. Kato Kalen was one of them, Kim Philby, the notorious British Spy was another, Kathy Webb who lied about being raped was another; all the great liars of CNN. In each of them the person is talking and there is a micro expression embedded, and afterwards the person is asked to rate the emotions they saw. And there is the training versus control group, it's a whopping effect. Now, we didn't get them perfect in 20 minutes. So the study that is now underway is doubling the training time, we think we will move them up quite a bit.

Emotional Skills

Paul Ekman
University of California at
San Francisco

There are also subtle expressions which are every bit as important but different. They are little tiny expressions. They are evident when emotion first begins, very often before the person knows how they are feeling. They are also signs of leakage as a person is deliberately concealing, and instead of getting a micro expression, they wiped everything off of their face expect a fragment, and that is what you see. Let me show you the subtle expression training tool, it is set up in a slightly different way. It uses only one person rather than 56 different people, the reason for that is because in order to see subtle expressions you have to hold the physical features constant. As it's coming up I will tell you that there is some evidence in terms of being able to voluntarily control your facial muscles and doing what I'm doing, which less than 5% of the population can do, but if one identical twin can so can the other, but not so for fraternal twins. For each emotion you get to see the subtle expressions and you get to pick the speed at which you will get to see it. And you work your way through and see these for each of the emotions. Then, when you're done, you go to practice and in practice, all of these subtle expressions from all the emotions come at you in a random order. Each time you take it it's a different random order and you choose the speed and you get the feedback. So I'll give you a reasonably fast speed and at the end you find out your percentage correct and people work at this, without being encouraged to, in order to get up to 100% which people get within an hour. So suppose you call that fear and you're wrong, it just gives it to you again. And then you click to the next one, and you go through them all and then you get your percentage correct.

Emotional Skills

Paul Ekman
University of California at
San Francisco

Let's just compare the micro expressions with the subtle expressions. In terms of the difference in speed, micros are always fast, whereas subtle may or may not be fast. Intensity of micro is extreme, in subtle it's slight. In micros it involves the whole face, in subtle one region. If you're interested in playing with these, either to get better or to try them, I put these out on the internet site in June and now there are seven teams working with them with Asbergers. There are four psychiatry departments using them in training. The Foreign Service Institute is using them in training all Foreign Service officers. So, they are getting used.

I want to turn now to the last of the four skills. Using the information about how others are feelings constructively. The METSET empowers you to get information that in Irving Goffman's terms, you weren't given. I know that Irving would say you're stealing information from people. So, it becomes important how you use that in a way that it's not going to make people feel you're invading their privacy, which indeed you are. And that you're doing it in a way that will be helpful rather than exploitive.

Emotional Skills

Paul Ekman
University of California at
San Francisco

First, never presume you know what's causing the emotion. Emotions don't tell you their source. You can pretty much guess from the context that Kato Kalen is scornful and angry at Marsha Clark, but maybe he's thinking about a fight with his wife that morning, or maybe he's angry at himself for starting a lie that now he got caught in. The first thing to find out is what is the emotion's source? Usually, it's within the social context, but don't be too egocentric and always presume, "I am the source." In my book, I did examples of what to consider in family life, in the workplace and in friendships, so they are very different considerations, I believe with different entitlements. I have a good enough relationship with my grown children that I feel entitled to say to them, "Did you have a bad day, did something that just happened annoy you? I don't feel entitled to that with my wife, I think it's a different relationship with your spouse than with your kids, but it depends on your spouse and it depends on your kids.

Emotional Skills

Paul Ekman
University of California at
San Francisco

For each emotion, I give three different contexts to consider in how you use this information. I just want to emphasize some general guidelines. Often the best course is to say nothing about what you've seen but just be alert to the possibilities, or you might say is there anything more you want to say about how you are feeling? Or, you might say I had the impression you were just feeling something more than what you told me about. You might be more specific, asking about the emotion you spotted. But how you respond depends on the nature of relationship; its past history and intended future and your knowledge of that person. You may not always be entitled to comment, even vaguely on the emotion you detected. Although, I don't believe that relationships work better when people understand and acknowledge how each other feel, that isn't always so.

As I emphasized earlier, some skills are hard to acquire and some are easy. The first two I described are hard to acquire and as best I can determine they require at least weekly practice, probably daily practice for maintenance. The Dalai Lama told me that if he is traveling and he is not able to mediate for three days he loses his impulse awareness. And incidentally, when I was a patient in psychoanalysis way back when I believed in that, I went four hours a week and I thought, "Boy that is a lot of time invested in behavior change." Four hours a week? These Buddhists do 70 hours a week, 80 hours a week, year after year after year. They are the Olympic athletics of the mind who have resculpted their brains.

Emotional Skills

Paul Ekman
University of California at
San Francisco

So, I've described four emotional skills. The first two I think are difficult to acquire and require daily practice. The next two I think are very easy to acquire and I think that you rapidly reach asymptote; you don't need a lot of practice. But skills without knowledge aren't enough to improve emotional life. You have to understand each emotion - its storyline, the universal themes that trigger that emotion, some of the more common variations on those themes, the appraisals that are commonly involved, the function of the emotion and what it does for us and how it relates to moods and when and how it becomes involved in emotional disorders. It's a tall order, those are all issues I addressed but I certainly haven't resolved in my book, but they are all things that we need to consider. Again, I want to emphasize that I've only described the four goals that I've either developed tools for or have been thinking a lot about. The others I think are all relevant, to be leading a skillful emotional life, one that is constructive to one's self and others. Thank you very much.

<div style="border:1px solid">

DISCUSSION

</div>

DISCUSSION

DR. ROBERTS: Thank you very much, questions?

DR. GADE: Do you find that each emotion is treated differently in terms of these four skills? For example, in my own introspection I think I'm pretty good at becoming aware when I'm becoming emotional in an angry way but probably not in a sad way.

DR. EKMAN: I have no data on that. Since individuals differ in almost everything that's important in their life, I suspect that there are individual differences, for example, in which emotions are most accessible to consciousness and which are not. But we haven't studied that, nobody has studied that that I know of. But if I were going to do research on that, which I'm not, that would be my bet. On the other hand, the Buddhist's report that these exercises and mindfulness is only one of them, benefits all emotions. But that's their report; I mean they report other things that I know are totally wrong, not because they want to be wrong. We all report our own experience which is in itself idiosyncratic. I don't know, it's a great question it could keep a whole laboratory busy for many years, if it were to be started.

DR. KYLLONEN: Has the method of lie detection that you talked about been used in the court system?

DISCUSSION

DR. EKMAN: I've only appeared as an expert witness once. This was a guy that allegedly murdered his wife, by riding back and forth over her in her own car. When they came to arrest him, he was reading my book, "Telling Lies." And so all they wanted me to do was to testify about what was in the book. I don't like how psychologists or psychiatrists participate in the judicial system in the United States, I think it's a very non-scientific way the jury can't find out how many other experts you've seen before this one, so I won't do it. I will work with law enforcement, I'll work giving them advice about interviews they have conducted or are about to conduct, and I do give training to law enforcement and judges. I have rewritten federal instructions to jury's about how to evaluate credibility, but I will not appear as an expert in a court unless we a get a system where the jury gets to know how many experts were consulted and what was their consensus. Which incidentally, occurs in the U.K; the U.K. is so far ahead of us in the criminal justice world it's shameful, for us, not for them.

DR. MATTHEWS: You talk about four different kinds of competencies or skills, but if people don't support some general intelligence they would have to be intercorrelated. So, do you have a sense to follow these different areas of competence to see if they are correlated or if they are independent?

DISCUSSION

DR. EKMAN: I think the first two are but I don't really know if they are correlated with the last two. This is really amazing to me, that something that is so valuable as recognizing subtle and micro expressions, people don't know. And even experienced policemen, don't know, and experienced lawyers; they just don't get it, they don't see it. But yet, they can learn it so fast. So, I'm delighted actually about both, because if they already knew it then what would I have to teach them? And if they couldn't learn it then again, what would I have to teach them? I've written an article called, "Why don't we catch liars?" It offers five different explanations about why we don't already have these skills which are acquirable. As I mentioned, the before and after studies are only just now starting to demonstrate this ...

DR. ZEIDNER: You seem to take issue with the emotional component of emotional intelligence but when I try to map your four skills into some of the current models, whether it be the Salovey, Mayer, or the recent Goleman orientation, which leaves sort of a four cell, two-dimensional space. But it's basically the same factors. The way you go about perhaps breaking them down or specifying their operationalization may differ; it may be a bit more constricted than some of the others, but basically they're also talking about identification and awareness of emotion in yourself and others. If you're basically using the same terms, why do you have problems with the emotional component of emotional intelligence?

DISCUSSION

DR. EKMAN: I would not claim expertise on the emotional intelligence literature, just a bit of knowledge of it. But my limited impression is they're talking about things that are more like enduring traits of how emotions are manifest in an interpersonal relationship, than with more momentary emotional states. Not that the two are unrelated, I think they are related. I wouldn't be willing to push that very far, maybe if this is relevant to emotional intelligence, good. It makes me feel more legitimate for being here.

DR. ZEIDNER: One other question has to do with the positive manifold you found across various channels in terms of latency. Could this be due to the negative affectivity factor?

DISCUSSION

DR. EKMAN: There are different signals and autonomic patterns I know people like John Cacioppo have a strong belief that that doesn't exist, but Levinson and I, and I could name you eight or nine other investigators, have strong multicultural evidence that there is a different autonomic signature for these different negative emotions. I don't think any emotion is negative, I think some emotional episodes are negative in a sense that they are destructive for the other person and for you. But, an enormous number of people enjoy being disgusted -just look at your kids! And you may think that sadness is a toxic emotion but look at the number of people who go to see movies and tell each other about novels that will make you cry. So, there is no question in my mind that the positive-negative distinction is a gloss that does harm to our understanding of the world of emotion and to actually how people behave; they don't behave in terms of negatives. The emotions have a lot of commonality but there are running in different musculatures, in different physiologies. Any cardiologist will tell you that individuals differ in their cardiac activities, and those are differences are apparent fairly early. So yes, you're going to find some consistency. For example, fear and anger are both heart rate accelerators, but the hands get hot in anger and cold in fear. So there are some commonalities but some differences, that's why we used a global autonomic measure to look at magnitude of response. I am a sharpener and not a lumper and I don't think we should lump it until after we have made the distinctions and gotten the evidence to show these distinctions don't get us anywhere. Rather, we should start with that kind approach, and you never find out what you're losing.

DISCUSSION

DR. LIVELY: I really like your example about Darwin's son. It seems as though there might be a contradiction in that and what you're saying about all the policemen and firemen and people who don't know how to recognize and supposedly then can't respond to emotions. So, I wonder if there is some developmental process?

DR. EKMAN: No, I think you missed an earlier point and I probably went over it too fast. When people are not trying to conceal or inhibit their emotions, you don't need to teach anybody, everybody gets it, we have evidence for that. What you need to teach people is to recognize subtle and micro expressions, without a doubt that is what most people miss. When people ask me what's the one book I should read I say Darwin's "Expression of Emotion" book. Even though I have 15 books out there, read Darwin's; it's the best book out there. But thank goodness, he left a few things alone, he only has two sentences in that book about deception; he left that for me. So, he didn't really consider the issue of concealment or deception.

DR. LIVELY: Do you think some people have it without knowing consciously they have it, and if so how do they get it?

DR. EKMAN: Maureen O'Sullivan is going to report on that after the break.

DISCUSSION

PARTICIPANT: I have a question regarding emotional profile. You say that trauma can move anyone into cell four, but do you know why people start out where they do? Do you think it's genetic, or is there something during people's childhood that cause people to start off being in any of these cells?

DR. EKMAN: Nobody has the answer to that question; I'm not the only one who can't give you the answer. These days, the rage and neuroscience is plasticity and I get into big arguments with my friend Richie Davidson, who is a big promoter of plasticity based on evidence largely from the motor system. Is there as much plasticity for emotion? I think we have to watch out that we don't become Lockeans again, which is what I think the neuroscientists are doing. So, we don't really know. In the whole temperament tradition, which really hasn't looked at specific emotions, says there is 50% inheritability. I think it is really the best evidence around. So, nothing is all the product of one thing, it's all interactive. Experience does change, but I believe in terms of the emotion system, there are constraints on changeability and that you have to really wallop people hard in order to change their profile. However, I can't conceive of a human subjects committee that would allow me to wallop them, even if I wanted to do the ecstasy thing I could never get that though my human subjects committee.
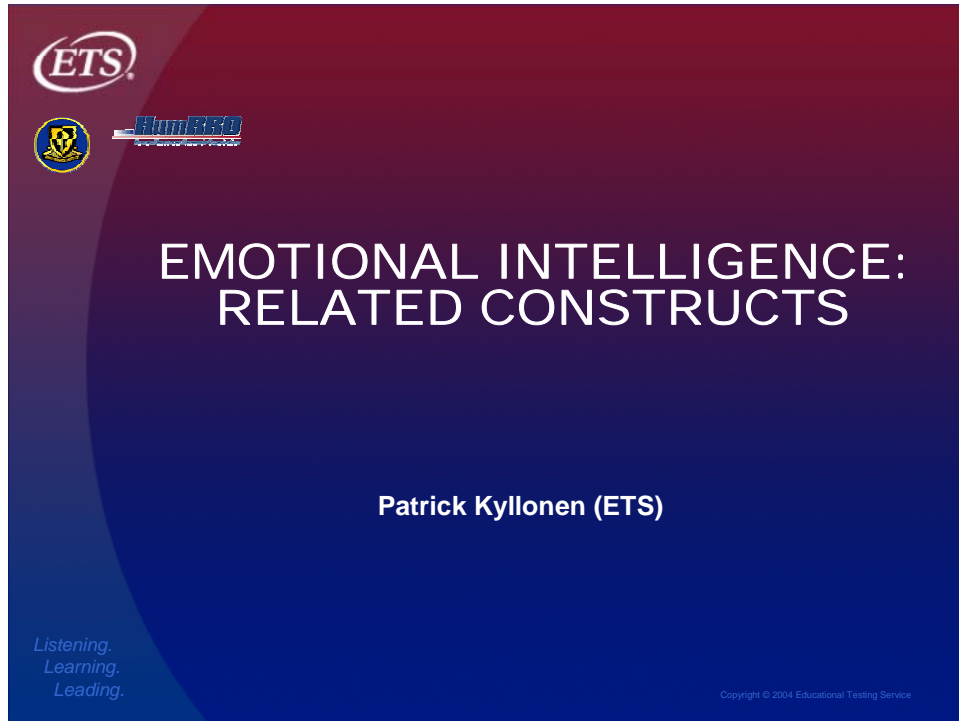
DR. KYLLONEN: I'm very excited about the next session. One of the strengths of this program is the diversity of perspectives and the diversity of approaches and findings, and we certainly are going to see that this morning. So what I would like to do is start off by introducing Professor Lawrence James from the University of Tennessee. He is going to talk about the Conditional Reasoning Approach to Personality Assessment.

The Conditional Reasoning Approach to
Personality Assessment

Lawrence R. James
The University of Tennessee

1

THE CONDITIONAL REASONING APPROACH TO PERSONALITY ASSESSMENT

DR. JAMES: Let me thank Richard for the opportunity to address this audience, and also for scheduling me right after a legend. I'm not really going to talk about emotions today and only tangentially about intelligence. Rather, I'd like to address the issue of measurement systems of personality and a new system we've been working with.

Objectives

1. Introduce a new system for personality measurement:
Conditional Reasoning Test for Aggression

2. Discuss how system enhances our ability to predict which
individuals will:
- engage in physical assault
- be verbally abusive
- quit, often in a disruptive manner (e.g., without notice)
- be habitually absent and/or tardy
- lie
- steal
- perform poorly
- be undependable

2

And the objectives are to introduce a new system for personality measurement, referred to as the Conditional Reasoning Test for Aggression. Aggression will be our personality construct of interest and I will discuss how this system enhances our ability to predict whether individuals will engage in: physical assault, verbal abuse, lying, stealing, or poor performance. Our hypothesis is that many acts of undependable behavior or deviance are products of indirect aggression or passive aggression intended to seek retribution for perceived injustice or victimization. My background is industrial psychology, so a good part of our work is done in the field in industrial situations.

Background

Everyday Reasoning:
Many waking hours are spent thinking inductively.

For example:
- What are the pros and cons of awarding prescription privileges to psychologists?
- What strategies might terrorists use in the future?
- Which is more conducive to education, public schools or private schools?
- What is the relative effect of diet on health?

3

I'd like to start with the basic idea of what is the type of reasoning that we are addressing. We spend many waking hours thinking inductively and we could ask, "What are the pros and cons of awarding prescription privileges to psychologists?" "What strategies might terrorists use in the future?" "Which is more conducive to education, public schools or private schools?" And, "What is the relative effect of diet on health?" One assesses which evidence is credible, which assumptions are tenable, which arguments are valid, and ultimately which conclusion or perhaps set of conclusions, which don't necessarily have to agree, might be true. This is inductive reasoning, although it focuses less on formal problem solving than on the deliberations that people use everyday to attempt to make sense of their environment or to decide upon reasonable ways to adapt to those environments.

A significant portion of everyday reasoning attempts to determine what a reasonable person would do when faced with frustrating or threatening circumstances.

- What would a reasonable person do upon learning that she or he has been passed over for promotion?
- How would a reasonable person react to being kept awake at night by the barking of a neighbor's dog?
- What is a reasonable response to being hit from behind while sitting in a car at a stoplight?
- What is a reasonable amount of time to wait in an examining room for a physician to appear?

4

A significant portion of everyday reasoning attempts to determine what a reasonable person would do when faced with frustrating or threatening circumstances. What would a reasonable person do upon learning that he or she has been passed over for promotion? Or, how would a reasonable person react to being kept awake at night by the barking of a neighbor's dog? Or, what is a reasonable amount of time to wait in an examining room for a physician to appear?

Individual Differences

People will differ in what they consider reasonable responses in each of these situations.

Reasoning of normal people is bounded by the cultural norms (i.e., standards) and cultural values (i.e., ideologies and rationales that support the norms) that define prosocial or socially adaptive behavior in our society.

5

Our concern is with individual differences; people will differ in what they consider to be reasonable responses in each of these situations. Basically, however, the inferences that pro-social people make about reasonable behavior are shaped by having been socialized into a culture, which is to say, having internalized the culture's norms, ideologies and rationales. Inference is consistent with cultural norms, whereas ideologies and rationales tend to be taken as reasonable, sound and sensible.

Justification of Aggressive Behavior

With this cognitive mindset, prosocial individuals typically find it difficult to understand why people engage in aggressive, antisocial behavior.
- Why would someone shoot their supervisor or neighbor (or neighbor's dog)?
- Why would someone steal major items from their company?
- Why would some students vandalize their own schools?
- Why do some people continuously engage in spreading malicious gossip?
- Why do some people try to control and dominate others?
- Why are some people habitually absent or tardy?
- How do some people justify deliberately inferior performance?

6

Let's then talk about justification of aggressive behavior. With this cognitive mindset, pro-social individuals typically find it difficult to understand why people engage in aggressive or antisocial behavior. Nothing in their cognitive repertoire provides them with a balance to explain or prepares them to answer questions like "Why would someone shoot their supervisor or their neighbor or their neighbor's dog?" Or, "Why would someone steal items from their company?" How do people justify deliberately inferior behavior? All the pro-social people can say about dysfunctional behaviors or social deviant behaviors is that they are socially deviant, according to the standards of the culture, and these people who participate or engage in these types of behaviors tend to be irrational or foolish.

How Aggressive People Think

Example: As a result of being passed over for promotion, the supervisor is physically attacked. People who attack their supervisors often regard their aggression as a justifiable act of retaliation for having been persecuted and victimized.

Chain of logic that leads to this justification typically begins with how supervisors are FRAMED.   Future aggressors frame interactions with their supervisors as dominance contests in which the bosses use their institutional power to force the subordinates into submissiveness.

Supervisors are thus framed as "intimidators" and "persecutors."

7

Now, pro-social people then oftentimes are mystified to learn that people who perpetrate aggressive behaviors often think of their actions as being reasonable. Now, how do aggressive people think? And let's use as an illustration here, a result of a person who has been passed over for promotion and subsequently attacks their supervisor. People who attack their supervisor often regard their aggression as a justifiable act of retaliation for having been persecuted or victimized. The chain of logic that leads to this justification typically begins with how the supervisor, who's been the target of aggression, is framed. Future aggressors tend to frame their interactions with supervisors as dominance contests, in which bosses use their institutional power to force the subordinates into submissiveness. Supervisors are thus framed as intimidators or perhaps ever persecutors. One of the things I want to do is to emphasize how important differences in framing are, and I'm not so much concerned about quantitative differences as I am concerned with qualitative differences.

Reasoning of Aggressive Person

The aggressive person reasons that confronting persecutors is more reasonable than cooperating with them because cooperation shows weakness and invites being harassed and bullied.

Confrontation, on the other hand, is thought to demonstrate bravery and forcefulness, which function to discourage supervisors from engaging in persecution.

8

So, the aggressive person reasons that confronting the persecutor is more reasonable than cooperating with them because cooperation shows weakness and invites being harassed and bullied. So, the basic idea is if you go along with the supervisor and conform to their wishes, then essentially what you're doing is reinforcing the behavior and you are likely to be bullied more in the future. Confrontation, on the other hand, is thought to demonstrate bravery and forcefulness, which functions to discourage the supervisor from engaging in persecution.

To justify confrontation, aggressive individuals often seek evidence of hostile intent on the part of their supervisors.

For example, they may attempt to build a case that their supervisors purposely overload them with work so that failing is inevitable.

9

To justify confrontation, aggressive individuals often seek evidence of hostile intent on the part of their supervisors. This would be seeking out information to support their beliefs or one could think of it as a confirmatory bias. For example, they may attempt to build a case that their supervisors purposely overload them with work, so failing is inevitable.

Not only does such reasoning imply a strong bias to attribute behavior to malevolent purpose and harmful intent, but also it suggests an implicit (i.e., unconscious) propensity to assume that the powerful will inflict harm on the less powerful.

This latent propensity maps into conscious reasoning by advancing an inference that one is the victim of inequity, exploitation, injustice, and oppression by those who are more powerful in one's life (e.g., supervisors as well as parents, teachers, employing organizations, or institutions such as the Internal Revenue Service).

10

Not only does such reasoning imply a strong bias to attribute behavior to malevolent purpose and harmful intent, but also it suggests an implicit, that is to say unconscious, propensity to assume that the powerful will inflict harm on the less powerful. This latent propensity maps into conscious reasoning by advancing an inference that one is the victim of inequity, exploitation, injustice, and oppression by those who are more powerful than in one's life. This might be supervisors as well as parents, teachers, employing organizations, or institutions such as the internal revenue service.

The thinking of aggressive individuals focuses progressively on how to obtain retribution and get even for the wounded pride, persecution, and disrespect that they believe they have suffered at the hands of their supervisors.

Some form of aggression is judged to be reasonable and justified because it offers a means to restore respect and to exact restitution for the perceived wrongs.

11

Thinking of aggressive people focuses progressively on how to obtain retribution or get even for the wounded pride, persecution, and disrespect that they believe that they have suffered at the hands of their supervisors. Some form of aggression is judged to be reasonable and justified because it offers a means to restore respect into exact restitution for the perceived wrongs.

Acts of retribution can take many forms, including

- verbal aggression (e.g., threats, character assassination, intimidation, spreading of malicious gossip),

- physical aggression (e.g., shoving, fighting, attacks with weapons), or

- passive aggression (e.g., intentional nonattendance at team meetings, deliberate failure to return messages; habitual absenteeism/tardiness).

12

Acts of retribution can take many forms, including verbal aggression, physical aggression or passive aggression.

The framing and reasoning of our illustrative aggressive persons differs considerably from the reasoning that would be expected by prosocial individuals

Aggressive Individuals

- supervisors are persecutors

- confrontation demonstrates power and discourages being victimized

- cooperation and compromise demonstrate weakness and invite attack

- retribution is justified

Prosocial Individuals

- supervisors are legitimate authority figures , perhaps even benefactors or protectors

- confrontation serves as a catalyst for overt and continued conflict

- cooperation and compromise defuse anger and promote long-term harmony

- nothing has occurred to justify retribution

13

Now, let's review just a little bit and add then the thinking of pro-social individuals. The framing and reasoning of our illustrative aggressive individuals differs considerably from the reasoning that would be expected from pro-social individuals. Whereas with aggressive individuals, supervisors are framed as persecutors, pro-social individuals would be much more prone to frame supervisors as legitimate authority figures perhaps even benefactors or protectors. While the aggressive person believes that confrontation demonstrates power and discourages being victimized, the pro-social person is likely to see confrontation as a catalyst for overt and continued conflict. Cooperation and compromise to the aggressive person demonstrates weakness and invites attack, whereas cooperation and compromise are seen by pro-social individuals as being used to defuse anger and promote long-term harmony. And the aggressive person believes that retribution is justified, whereas the pro-social person does not. Basically, aggressive individuals differ from pro-social individuals in terms of the adjectives they use to frame persons in authority, the behaviors of their supervisors that they select as being relevant indicators of their supervisor's intentions, the assumptions they make about why supervisors behave as they do, and the arguments for and against aggression that they judge to be valid. Let me emphasize again that a large number of these differences are based on qualitative differences not just quantitative differences.

Why do aggressive people frame and reason as they do?

The framing and reasoning of aggressive individuals is designed to justify a strong motive to harm others. Framing and reasoning are intended to enhance the rational appeal of engaging in behaviors that harm others. This is rationalization or justification.

Aggressive people are generally unaware of this intent. They believe their reasoning is rational.

Aggressive people can engage in this self-deception because the biasing of their reasoning (toward rationalizing aggression) occurs below the surface of their consciousness.

14

Why do aggressive people frame and reason as they do? Is the framing and reasoning of aggressive individuals designed to justify a strong motive to harm others? Framing and reasoning are intended to enhance the rational appeal of engaging in behaviors that harm others. This is rationalization or justification. Aggressive people are generally unaware of this intent; they believe that their reasoning is rational. Aggressive people can engage in this self-deception because the biasing of their reasoning, toward rationalizing aggression, occurs below the surface of their consciousness. I'm trained basically as a psychometrician, the use of terms like "unconscious" and "implicit" is tough, but we're getting there.

Specifically, the reasoning of aggressive people is implicitly (unconsciously) shaped by biases whose function is to enhance the rational appeal of aggression.

James (1998; James & Mazerolle, 2002) proposed the term "justification mechanisms" to identify these implicit biases.

15

Specifically, the reasoning of aggressive people is implicitly shaped by biases whose function is to enhance the rational appeal of aggression. We propose that one could use the term "justification mechanisms" to describe these particular implicit biases.

Justification Mechanisms (JMs)

(1) The Potency Bias involves an implicit proclivity to frame interactions with others as contests to establish dominance versus submissiveness.

16

There are six justification mechanisms that we have identified in the literature that tend to characterize the reasoning of aggressive individuals. Potency Bias involves an implicit proclivity and can frame interactions with others as contests to establish dominance versus submissiveness. The basic idea is that people tend to perceive through a lens that judges behaviors of others in terms of its dominance versus submissiveness. Such framing is the cornerstone for justifying aggression as a demonstration of strength, bravery, or fearlessness, and not acting aggressively is logically attributed to weakness, fear, cowardice, or impudence.

JMs

(2) The Hostile Attribution Bias consists of an implicit predilection to assume that malevolent purpose or harmful intent is the primary motivation underlying the behaviors of others (see Crick & Dodge, 1994).

17

An aggressive person may thus rationalize aggression by reasoning that aggression is an act of strength or bravery that gains respect from others, or to show weaknesses to invite powerful others to take advantage of you. This is the Hostile Attribution Bias; it's well known in the literature and consists of an implicit predilection to assume that malevolent purpose or harmful intent is the primary motivation underlying the behaviors of others. Basically, the actions of others pass through a perceptual prism primed to see hostile or malevolent intent. The attributions of hostile intent that follow are then employed by the aggressive person to rationalize his or her own hostile behaviors as acts of self defense intended to ward off physical or verbal attack.

JMs

(3) The Retribution Bias involves an unconscious tendency to confer logical priority to reparation or retaliation over reconciliation.

18

The Retribution Bias involves an unconscious tendency to confer logical priority to reparation or retaliation over reconciliation. This justification mechanism promotes reasoning that aggression is a justifiable response, if the intent of the response is to restore, respect or do exact restitution for a perceived wrong. Retaliation thus appears more reasonable than forgiveness, vindication appears more reasonable than cooperation, and obtaining revenge appears more reasonable than maintaining the relationship. This bias often underlies rationalizations for aggression engendered by wounded pride, challenged self-esteem, or perceived disrespect.

JMs
(4) The Victimization by Powerful Others Bias is an implicit tendency to see everyday people (including oneself) as victims of inequity, exploitation, injustice, or oppression by powerful others.

19

The Victimization by Powerful Others Bias is an implicit tendency to see everyday people, including oneself, as victims of inequity, exploitation, injustice, or oppression by powerful others. Powerful others include entities such as supervisors, teachers, and employment organizations. The framing of events hypothesis about cause and effect and confirmatory searches for evidence, both engender and reinforce unconscious inclinations to infer that people are being victimized by powerful others. This reasoning furnishes the foundation for rationalizing acts of aggression as warranted corrections of inequities or legitimate strikes against oppression.

JMs

(5) Derogation of Target Bias consists of an unconscious tendency to characterize those one wishes to make (or has made) targets of aggression as evil, immoral, or untrustworthy.

20

The Derogation of Target Bias, and this been the most difficult one for us to make work, consists of an unconscious tendency to characterize those one wishes to make, or has made, targets of aggression as evil, immoral or untrustworthy. And to infer or associate such negative traits with the target makes the target more deserving of aggression.

JMs

(6) Social Discounting Bias is comprised by an implicit tendency to favor socially unorthodox and antisocial reasons as logically probable causes of social events and relationships.

21

And finally, the Social Discounting Bias, and out of the group this analytically has shown to be the most important, is comprised by an implicit tendency to favor socially unorthodox and antisocial reasons as logically probable causes of social events and relationships. Reasoning shaped by this bias reflects disdain for traditional ideas and intentional beliefs. This reasoning will further evidence a lack of sensitivity, empathy, and concern for social customs, often accompanied by the absence of rational prohibitions against acting aggressively. Analyses of social events tend toward the cynical and critical with the proclivity to associate positive outcomes with aggression.

Measuring JMs

How is it possible to determine whether JMs are instrumental in shaping a person's reasoning?

Why would we want to know this?

22

Now, the basic idea of the measurement system is to measure these justification mechanisms. How is it possible to determine whether justification mechanisms are instrumental in shaping a person's reasoning? And why would we want to know this?

Measuring JMs

Reasoning that is shaped by JMs indicates a predisposition to behave aggressively. The person is prepared to rationalize aggression.

23

Reasoning that is shaped by justification mechanisms indicates a predisposition to behave aggressively. This person is prepared to rationalize aggression and to engage in aggression. We are working basically from the assumption that only people who have aggressed in the past and intend to aggress in the future have needs of justification mechanisms for aggression. We can use the measurements of this, if we can obtain them, as a basis for diagnosing aggressive tendencies or as a foundation for scientific studies of implicit cognitive process.

Conditional Reasoning

Illustrative problem for entry-level jobs:

The old saying, "an eye for an eye," means that if someone hurts you, then you should hurt that person back. If you are hit, then you should hit back. If someone burns your house, then you should burn that person's house.

Which of the following is the biggest problem with the "eye for an eye" plan?

a. It tells people to "turn the other cheek."

b. It offers no way to settle a conflict in a friendly manner.

c. It can only be used at certain times of the year.

d. People have to wait until they are attacked before they can strike.

24

This is a problem we've designed for entry level jobs for students.

Respondents believe that their critical intellectual skills guide their attempts to identify a logically correct conclusion to this problem.

The demand appears intellectual because inductive reasoning is in fact required to solve the problem.

However, accuracy of inductive analyses shaped by critical reasoning skills is NOT the primary issue here.

25

Respondents believe that their critical intellectual skills guide their attempts to identify a logically correct conclusion to the problem. The people are given the problems and they are told that it's a reasoning test. Demand appears intellectual because inductive reasoning is required to solve the problem. However, accuracy of inductive analyses shaped by critical reasoning skills is not the primary issue here.

Our intent is that inductive seeking of the most reasonable answer be guided, unknowingly, by implicit assumptions about what constitutes rational behavior.

We want reasoning to be shaped by either JMs for aggression or internalized prosocial values and ideologies.

In a manner analogous to everyday reasoning about the rationality of aggression, what is judged to be the most reasonable answer to the problem is CONDITIONAL on whether analyses are shaped by JMs or prosocial proclivities.

26

Our intent is that inductive seeking is the most reasonable answer to be guided unknowingly by implicit assumptions about what constitutes rational behavior. We want reasoning to be shaped by either justification mechanisms for aggression or internalized pro-social values and ideologies. The idea here is that the pro-social values and ideologies are largely automated. In a manner of analogous everyday reasoning about the rationality of aggression, what is judged to be the most reasonable answer to a problem is conditional on whether analyses are shaped by justification mechanisms or pro-social proclivities.

d. People have to wait until they are attacked before they can strike.

This alternative was designed to be logically attractive to aggressive individuals because it is grounded in their biases. Specifically, it

- tacitly promotes retribution as being logically preferable to reconciliation (Retribution Bias)

- is founded on the unstated assumption that the powerful will inflict harm on the less powerful unless the less powerful strike first (Victimization & Potency Biases)

27

For example, people have to wait until they are attacked before they can strike. This alternative was designed to be logically attractive to aggressive individuals because it's grounded in their biases. Specifically, it tacitly promotes retribution as being logically preferable to reconciliation; we use the Retribution Bias to shape that one. And furthermore, is founded on the unstated assumption that the powerful will inflict harm on the less powerful unless the less powerful strikes first. And that is based on the Victimization and Potency Biases.

The logical credibility of this answer increases as JMs become more prominent in shaping reasoning.

Selection of Alternative d as the most logically persuasive of the conclusions offered provides indirect evidence that the Retribution, Victimization, and Potency Biases are implicitly instrumental in shaping the reasoning of a respondent.

28

Now, the logical credibility of this answer increases as justification mechanisms become more prominent in shaping reasoning. And this is another way of looking at conditional reasoning, which is that the logical credibility of an answer is conditional on the prominence of justification mechanisms in shaping reasoning. Selection of alternative "d" as the most logically persuasive of the conclusions offered provides indirect evidence that the Retribution, Victimization and Potency Biases are implicitly instrumental in shaping the reasoning of a respondent.

b. It offers no way to settle a conflict in a friendly manner.

This alternative was designed to be logically attractive to prosocial individuals because it
- promotes a prosocial counterbalance to the antagonistic and provocative tenor of the aggression alternative, and
- is grounded in the unstated assumption that conflict is logically less reasonable than compromise and cooperation.

29

Let's go to the second alternative. The eye-for-an-eye offers no way to settle a conflict in a friendly manner, this alternative was designed to be logically attractive to pro-social individuals because it promotes a pro-social counterbalance to the antagonistic and provocative tenor of the aggression alternative. Furthermore, it is grounded in the unstated assumption that conflict is logically less reasonable than compromise and cooperation.

The logical plausibility of this answer increases as prosocial factors become more influential in shaping reasoning.

Selection of Alternative b provides indirect evidence that the Retribution, Victimization, and Potency Biases are NOT instrumental in shaping reasoning.

30

The logical plausibility of this answer increases as pro-social factors become more influential in shaping reasoning. So, selection of this alternative is indirect evidence that the Retribution, Victimization, and Potency biases are not instrumental in shaping the reasoning of response.

a. It tells people to "turn the other cheek."
c. It can only be used at certain times of the year.

These two alternatives are included to enhance the face validity of the task and to protect the indirect nature of measurement.

Alternatives a and c are meant to be clearly illogical and rejected by respondents (which is usually the case).

31

Now, the last two alternatives are included to enhance the face validity of the task and to protect the indirect nature of measurement. These alternatives are meant to be clearly illogical and rejected by respondents, which is usually the case; very few people ever pick these distracters. I debated whether to show you some of our new items. We find that respondents have no troubles with this, but psychologists do once in a while. So, what I decided to do primarily to enhance the sophistication of these problems is to take these distracters and make them real distracters. I spent a few months early in the game doing a content analysis of inductive reasoning problems and it became pretty clear that the trick to writing a really good inductive reasoning problem, well we'll see if people here will disagree with me, is writing good distracters and writing good distracters is tough business. So, what we've done is we've crossed justification mechanisms with intelligence. Basically, we still have two distracters - they are both wrong, but one of them is logically wrong and a measure of a justification mechanism and one of them is logically wrong and a measure of pro-social values. So it's a fully crossed model and there are much more sophisticated items. I don't have much data on them yet, and I have no validity so I am going to wait until I present those but we are working on some more sophisticated techniques.

How this illustrative problem is solved is dependent, or CONDITIONAL, on the relative degrees to which JMs for aggression versus prosocial values and ideologies are instrumental in shaping reasoning.

The problem is referred to as a CONDITIONAL REASONING PROBLEM.

This psychometric approach is referred to as the "conditional reasoning measurement system" (James, 1998; James & Mazarolle, 2002)

32

How this illustrative problem is solved is dependent or conditional on the relative degrees to which justification mechanisms for aggression versus pro-social values and ideologies are instrumental in shaping reasoning. The problem is referred to as a conditional reasoning problem. The psychometric approach is referred to as the conditional reasoning measurement system.

Animals in the wild often fight to determine who will breed. The reason for fighting, sometimes to the death, is that only the strongest will be allowed to breed. Fighting is designed to insure that only the most powerful and fit animals survive and reproduce.

Humans usually reproduce without first determining who is the strongest. This suggests that:

a. physical strength is usually not important in determining who will contribute to society.

b. animals breed most often in the Fall.

c. the study of genetics is becoming less popular.

d. the human race, on the average, is becoming physically weaker. [33]

I have a couple more problems and I'll let you play with these for a second. Which one is the aggression response? D, right. And the pro-social response? A, right. I've learned the hard way not to ask people how they got to this because they might tell you.

JMs

1. **Potency Bias**: a propensity to frame interactions with others through a prism of dominance versus submissiveness (or strength versus weakness).

2. **Hostile Attribution Bias**: a tendency to seek out malevolent purpose or harmful intent as the motivations for the actions of others.

3. **Retribution Bias**: a predilection to favor vengeance, retribution, and retaliation over reconciliation, cooperation, or compromise.

4. **Derogation of Target Bias**: a predilection to characterize the target of aggression as evil, immoral, untrustworthy, or corrupt, and thus more deserving of being assailed.

5. **Victimization by Powerful Others Bias:** tendency to frame self as a victim and to see self as being exploited and taken advantage of by the powerful (e.g., employing organization).

6. **Social Discounting Bias**: tendency to call on socially unorthodox and frequently antisocial beliefs to interpret and to analyze social events and relationships.

34

So, which justification mechanism bias for the aggression response? Potency, right.

One-half of marriages end in divorce. One reason for the large number of divorces is that it is easy and quick to get a divorce. If a couple can agree on how to split their property fairly, then they can get a divorce simply by filling out forms and taking them to court. They do not need lawyers.

This means that:

a. people are older when they get married.

b. if one's husband or wife hires a lawyer, then he or she is not planning to play fair.

c. couples might get back together if getting a divorce took longer.

d. more men than women get divorced.

35

One more. The aggression response? D, right. And the pro-social response? C, right.

JMs

1. **Potency Bias**: a propensity to frame interactions with others through a prism of dominance versus submissiveness (or strength versus weakness).

2. **Hostile Attribution Bias**: a tendency to seek out malevolent purpose or harmful intent as the motivations for the actions of others.

3. **Retribution Bias**: a predilection to favor vengeance, retribution, and retaliation over reconciliation, cooperation, or compromise.

4. **Derogation of Target Bias**: a predilection to characterize the target of aggression as evil, immoral, untrustworthy, or corrupt, and thus more deserving of being assailed.

5. **Victimization by Powerful Others Bias:** tendency to frame self as a victim and to see self as being exploited and taken advantage of by the powerful (e.g., employing organization).

6. **Social Discounting Bias**: tendency to call on socially unorthodox and frequently antisocial beliefs to interpret and to analyze social events and relationships.

36

And the justification mechanism?

Conditional Reasoning Test for Aggression (CRT-A)

Composition: 22 conditional reasoning (CR) problems and three inductive reasoning problems.

Scoring: "+1" for aggression (AG) alternative; "0" for every logically incorrect alternative; and "-1" for prosocial or nonaggressive (NA) alternative. Scores are summed to furnish a composite score on the "Justification of Aggression Scale" or "JAGS."

Interpretation of scores on the JAGS:

High score: JMs for aggression are instrumental in guiding and shaping reasoning. Respondents are "justifiers," which connotes that they are implicitly prepared and willing to engage in some form of aggressive behavior.

Low score: JMs are <u>not</u> instrumental in guiding and shaping reasoning. Respondents are not implicitly prepared and willing to engage in aggressive behavior.

Mid-range score: JMs are only sporadically instrumental in shaping reasoning. The implicit readiness to aggress is therefore likely to be only modest.

37

The instrument itself is composed of 22 problems, including the three you have just seen, and we have used a number of different scoring methods. It changes, but one point for an aggression response, zero for logically incorrect alternative, minus one for pro-social, this gives us the cleanest factor structure, the most straightforward way to simply score the number of aggression responses or number of pro-social depends on which way you want to orient the scale. We refer to this as the Justification of Aggression Scale or JAGS. The interpretation of the scores is that a high score reveals that justification mechanisms for aggression are instrumental in guiding and shaping and reasoning. Respondents are justifiers, which connotes that they are implicitly prepared and willing to engage in some form of aggressive behavior. A low score reveals that justification mechanisms are not instrumental in guiding and shaping and reasoning. Respondents are not implicitly prepared or willing to engage in aggressive behavior.

Verbal/Visual Conditional Reasoning Test (VCRT).

For less adept readers, a test based on a "verbal-visual" version of a subset of the CR problems was designed to have a threshold reading level of approximately the fifth to sixth grade (Green & James, 1999).

Referred to as the "VCRT," this test consists of bare-bones versions of CR problems. The problems are presented both verbally and in written form using a videocassette player and television.

The written component consists of simplified prose, which is overlaid on a photograph consistent with the basic theme of the CR problem.

The current VCRT contains 14 CR problems, 12 of which are shared with the CRT-A. Work continues on converting CRT-A problems to the VCRT format. Scoring and interpretation of scores on the VCRT are analogous to procedures employed on the JAGS.

38

We did this for an unnamed Southeastern University Football team; it's a Verbal/visual Conditional Reasoning Test, VCRT. Basically, what we do is simplify the problems or take it to three alternatives: an aggression response, a pro-social response, and one distracter and then overlay it on a photograph that was searched through the files of Life Magazine to find photographs that appeared to best illustrate this particular problem. It is also is given verbally; the problem is written in a simplified version and the problem is given verbally. So far we have 14 problems in this instrument, 12 of which are shared with the Conditional Reasoning Test for aggression.

Empirical Evaluations

Distributions on CR problems and JAGS were skewed. Approximately 12% of respondents considered moderate to strong justifiers.

CR problems factored by the JMs.

39

Empirical evidence arises from the fact that distributions on the CR problems and JAGS were skewed. Approximately 12 percent of respondents are considered moderate to strong justifiers or moderately strong aggressive. Additionally, the factor analyses demonstrate that the problems do indeed factor by justification mechanism, although there is a fair degree of complexity; eight particular items did not necessarily pick up just one justification mechanism.

Reliabilities

Factorial (n = 1603)
Factor 1   .87
Factor 2   .82
Factor 3   .81
Factor 4   .76
Factor 5   .74

Internal Consistency
CRT   .76 (n = 1603)
VCRT   .78 (n = 225)

Alternative Form
CRT/VCRT   .82 (n = 276)   40

   If you look at the factorial reliabilities, they tend to run fairly high. The internal consistency estimates tend to run in the mid seventies, sometimes the upper seventies. I plan to build a sufficient number of items that we have fairly internally consistent measures of each justification mechanism, but we are not there yet, we don't have enough problems. With an alternative form, where we gave the conditional reasoning test right at the first of the semester to a group of students who then took the VCRT toward the end of the semester, we calculated the correlation between them, which was .82, which demonstrates that the two ways of measuring are pretty similar as well as the stability, so that is pretty good.

Uncorrected Validities for Scores on JAGS

Sample # & Criterion	Sample	Instrument	Design	Uncorrected Validity*
1. Performance Ratings	140 Patrol Officers	CRT	Concurrent	-.49
2. Absences from Class	188 Undergrads	CRT	Predictive	.37
3. Lack of Truthfulness about Extra Credit	60 Undergrads	VCRT	Experiment	.49
4. Absences from Work	97 Nuclear Facility Operators	CRT	Postdictive	.42
5. Conduct Violations	225 Undergrads	VCRT	Postdictive	.55
6. Attrition	135 Restaurant Employees	CRT	Predictive	.32
7. Absences from Work	105 Package Handlers	CRT-A	Predictive	.34
8. Work Unreliability	111 Temporary Employees	CRT-A	Predictive	.43
9. Theft	95 Undergrads	CRT-A	Experiment	.64
10. Hard Fouls & Fights in Intramural Basketball	191 Undergrads	CRT-A	Predictive	.38

*All correlations are statistically significant (p<.05).

41

We conducted a number of validity studies, most with real people in real world situations. For example, patrol officers, nuclear facility operators, restaurant employees, package handlers, temporary employees, etc. We've also conducted a couple of experiments. In study three, we tested to see if the students would be truthful about the amount of extra credit they deserved in a particular exercise. In another study that was done by Terri Mitchell from the University of Washington, is whether students would steal a gift after they had been clearly frustrated in an exercise. What we found was that a number of people had questions about whether these criteria are actually measures of aggression. And so in response to that, we did a review of the literature and found basically aggression is an integral component of organizational retribution, counter-productive performance, dysfunctional resistance tactics, obstructive behaviors and workplace deviants. You can see where the research has been going the last ten years in organizational psychology. This is because many, if not most deviant retaliatory, counter productive, resistant and destructive and displaced behaviors, involve hostile attempts to harm an organization or its constituents by exacting retribution, revenge or retaliation in ways that disrupt work schedules, impede productivity, weaken morale, undermine authority, encourage rebelliousness and get even with the boss or co-workers.

Uncorrected Validities for Scores on JAGS

Sample # & Criterion	Sample	Instrument	Design	Uncorrected Validity*
1. Performance Ratings	140 Patrol Officers	CRT	Concurrent	-.49
2. Absences from Class	188 Undergrads	CRT	Predictive	.37
3. Lack of Truthfulness about Extra Credit	60 Undergrads	VCRT	Experiment	.49
4. Absences from Work	97 Nuclear Facility Operators	CRT	Postdictive	.42
5. Conduct Violations	225 Undergrads	VCRT	Postdictive	.55
6. Attrition	135 Restaurant Employees	CRT	Predictive	.32
7. Absences from Work	105 Package Handlers	CRT-A	Predictive	.34
8. Work Unreliability	111 Temporary Employees	CRT-A	Predictive	.43
9. Theft	95 Undergrads	CRT-A	Experiment	.64
10. Hard Fouls & Fights in Intramural Basketball	191 Undergrads	CRT-A	Predictive	.38

*All correlations are statistically significant (p<.05).

41

To avoid punishment, these processes seldom involve outright violence or acts that are easily detectible as aggression. Rather, they focus on indirect passive-aggressive behaviors such as failing to come to work, or coming to work late, stealing from those seen as guilty of injustices, lying to authority figures to regain face and obtain retribution for being disrespectful in performing in poor unreliable or improper matters. In short, we got a little defensive about this and said look, a lot of these have aggressive types of connotations or can be products of aggressive behavior, doesn't mean that all of them are, but the probability is pretty good. After all that we went ahead and did an aggression study. So, we studied intramural basketball at a large Southeastern University, and we looked at a lot of things. For example, we looked at hard fouls and fights in intramural basketball, and basically a hard foul is where someone is hit and often times goes to the ground or a fight. We had people watch every one of the games and recorded this and the students before the start of the season, took the conditional reasoning test.

Uncorrected Validities for Scores on JAGS

Sample # & Criterion	Sample	Instrument	Design	Uncorrected Validity*
1. Performance Ratings	140 Patrol Officers	CRT	Concurrent	-.49
2. Absences from Class	188 Undergrads	CRT	Predictive	.37
3. Lack of Truthfulness about Extra Credit	60 Undergrads	VCRT	Experiment	.49
4. Absences from Work	97 Nuclear Facility Operators	CRT	Postdictive	.42
5. Conduct Violations	225 Undergrads	VCRT	Postdictive	.55
6. Attrition	135 Restaurant Employees	CRT	Predictive	.32
7. Absences from Work	105 Package Handlers	CRT-A	Predictive	.34
8. Work Unreliability	111 Temporary Employees	CRT-A	Predictive	.43
9. Theft	95 Undergrads	CRT-A	Experiment	.64
10. Hard Fouls & Fights in Intramural Basketball	191 Undergrads	CRT-A	Predictive	.38

*All correlations are statistically significant (p<.05).

41

You take a look at the validities in this table you will see that the average validity runs about .44. Putting that in a context, you take a look at somewhere around a dozen or so meta-analyses that have been done in the last 10-12 years on validity and personality in industrial situations, almost inevitably based on self-report, the average validity is .12. So, you are looking at the difference and these are uncorrected validities, so you're looking at a difference of an average of .44 against .12 and I think if anything that makes a pretty good case that implicit personality has a place in psychology.

Correlations between Scores on the JAGS
and Critical Intellectual Skills

Study	Intelligence Scale	n	Sample	Correlation
2	ACT	188	Undergrads	-.06
3	ACT	60	Undergrads	-.05
5	ACT	225	Undergrads	-.08

*p < .05
Note. ACT = American College Testing.

42

There is no correlation between the JAGS and critical intellectual skills. If I were picking the test this school used, I would have used the SAT, but we have the ACT here, so sorry about that. Basically, how smart you are doesn't determine how you respond to the questions.

Relationships between Scores on the JAGS and Gender

Sample	Composition	% Female	Point Biserial Correlation	Biserial Correlation
2	188 undergrads	.34	-.08	-.10
3	60 Undergrads	.60	-.22*	-.29*
5	225 Undergrads	.49	-.20*	-.25*
6	120 Restaurant Employees	.66	-.06	-.08
7	105 Package Handlers	.26	.04	.05
8	111 Temporary Employees	.36	.07	.09
9	95 Undergrads	.55	.00	.00

Male = 0, Female = 1.
*p < .05

43

In terms of relationships between the scores on the Conditional Reasoning Test and gender, in a few cases we find that males tend to score a little more highly than females but that's certainly not consistent. Our basic position on this is that males and females have justification mechanisms and that the primary difference in manifestations of aggression comes in overtly; how it's manifested is not necessarily in the implicit possession of rationalization mechanisms.

Relationships between Scores on the JAGS and Race

Sample	Composition	%	Race	Relationship
2	188 undergrads	.90	White (0)	$r = .06^{1}$
		.10	African American (1)	$(r = .10)^{2}$
7	105 Package Handlers	.23	White	$F = 1.43$
		.52	African American	
		.24	Hispanic	
		.01	Asian	
8	111 Temporary Employees	.82	White (0)	$r = .07^{1}$
		.18	African American (1)	$(r = .10)^{2}$
9	95 Undergrads	.57	White	$F = .16$
		.38	Asian	
		.04	Hispanic	
		.01	African American	

$*p < .05$
[1]Point Biserial Correlation
[2]Biserial Correlation

44

There is no correlation between the JAGS and race.

Correlations between Scores on the JAGS and Self-Report Measures of Aggression			
Sample	Composition	Alternative Measure	Correlation
3	60 Undergrads	PRF Aggression	.14
		PRF Dominance	.05
		PRF Impulsivity	.11
5	225 Undergrads	NEO-PI-R Angry Hostility	.26*
		NEO-PI-R Dutifulness	-.18*
9	95 Undergrads	Aggression Questionnaire	.24*
10	191 Undergrads	NEO-PI-R Angry Hostility	.002

*p <. 05
PRF = Personality Research Form (Jackson, 1967). NEO-PI-R = NEO Personality Inventory (Revised) (Costa & McCrae, 1985, 1992). Aggression Questionnaire (Buss & Perry, 1992). 45

This is the one that always gets me in trouble, which is the correlation between the scores on the JAGS and self-report measures of aggression. For the most part, the correlation is either non-significant or significantly low. When I teach a personality class and when we get to this point right here and the industrial psychologists look at this and they say "Aha, lack of convergent validity, you're in trouble with construct validity, your implicit correlates with your explicit and you have problems." Then the clinical psychologists look at this and say, "What else is new? There obviously is a conflict between implicit and explicit, why would you expect there to be a correlation?" And what's interesting is that from that day forward, the clinical psychologists sit on one side of the room and the industrial psychologists sit on the other side. We are conducting a number of studies now, looking at the interaction between implicit and explicit using integrative models that McClellan proposed. We are finding some fairly good success with those. Basically, what you do is cross self-reports of aggression with the conditional reasoning measure of aggression and then you look for consistencies and inconsistencies. And what we're finding is that the most aggressive people are people who on the self-report measure say that they're aggressive and with our measures we say yes, you have the rationalization to back them up. But, we are finding also that some of the least aggressive with overt measures are people who say they're not aggressive but have the justification mechanisms in place. Unless we go to a passive-aggressive measure, then things change. We intend to get there but we are not there yet. This is a fascinating area and we will keep with it.

Multiple Regression-Dominance Analyses

Sample	Criterion	Multiple R	Variable	Zero-Order r	Beta Weight	Relative Importance
3	Lack of Truthfulness about Extra Credit	.55*	JAGS	.49*	.51*	82.83%
			ACT	-.07	-.12	2.89%
			PRF Aggression	.16	.17	8.11%
			PRF Impulsiveness	.14	.14	6.05%
			PRF Dominance	.05	.04	.72%
5	Student Conduct Violations	.61*	JAGS	.55*	.53*	77.60%
			NEO Angry Hostility	.26*	.21*	16.00%
			NEO Dutifulness	-.18*	-.13*	6.43%
9	Theft	.70*	JAGS	.64*	.57*	71.7%
			Aggression Quest.	.44*	.30*	28.3%
10	Hard Fouls & Fights	.45*	JAGS	.38*	.38*	74%
			NEO Angry Hostility	.22*	.22*	26%

*p <. 05

46

We are back at the old question, implicit versus explicit. I don't really like to frame it this way, and sometime as soon as I get some slides prepared on these channeling models, I'm going to get rid of this one. But, the basic idea was if we put these things into straight multiple regressions, now these don't have the interaction terms in them, and we compute a dominance analysis which takes a look at the relative importance of each of the predictors to the multiple R squared, you will see that the JAGS or the Conditional Reasoning measure wins out every time. I consider this important because for years as I did research on organizational and psychological climate, I became convinced that the most salient feature in terms of perceptions of organizational situations and organizational environment is who is doing the proceedings. Then, to relate that, it doesn't mean that the organization is not important, it just means that there is a strong interaction with who is doing the proceeding. Then the next step was then to relate the climate perceptions to personality factors to find out what it was about people that were leading to differences in perceptions of the environment. These were fairly complex models with person-environment fit and the further we got into this the more convinced I became that maybe the hypotheses weren't so bad, but it was the measurement system and maybe on both sides of the equation. The self- report measures weren't sophisticated enough to pick up some of the hypotheses we had and that is what got me into this venture. Basically, you can't walk away from meta-analyses that say that personality counts for one-percent of the variance in people's behavior and in organizations; it's much more important than that.

Closing Comments

The JAGS predicts a reasonably wide array of behavioral criteria, many of which have been salient to civilian, government, and military organizations for some time (e.g., attrition).

The average validity of .44 compares favorably with validities produced by measures of critical intellectual skills, and surpasses the average, uncorrected validity of .12 produced by self-report personality inventories.

Due to their indirect nature, it is reasonable to expect that CR problems are resilient to faking. Initial findings are that the CRTs for aggression and achievement motivation are resistant to socially desirable responding and impression management when administered under normal circumstances.

47

The JAGS predicts a reasonably wide and array of behavioral criteria, many of which have been salient to civilian, government, and military organizations for some time. Due to their indirect nature, it is reasonable to expect that Conditional Reasoning problems are resilient to faking. Initial findings are that the CRTs for aggression and achievement motivation are resistant to socially desirable responding and impression management when administered under normal circumstances.

DISCUSSION

DR. JAMES: Thank you.

DR. KYLLONEN: We have some time for questions.

DR. MAYER: Is there ever a time when these justification mechanisms might need to be adjusted?

DR. JAMES: Well, I am not going to make a moral judgment, but yeah, sure. Basically it is one rationalizing aggression. The way I prefer to look at this, is say that you were a member of the French resistance in Paris, and you engaged in aggression like sabotage, is that aggression? One can get into debates on that sort of thing, but the idea is that with aggressive people, they take that type of situation, which you have a justifiable form of aggression and try to apply it to situations in which it is not justified. That is the key to what we are trying to look at. Would there be situations in which the aggression is justified, therefore the rationalization, in which case it might not even be a rationalization? Sure!

DR. GRANDEY: Can we teach people who have problems like feeling, and wrong values about these things and have them and change what they believe?

DR. JAMES: I'm in the measurement business; I'm not in the change business. I suspect you could.

DISCUSSION

PARTICIPANT: Would you comment on the relationship of what you've been assessing with psychopathy?

DR. JAMES: Well Bob Hogan tells me it's pretty high, so I would suspect with the implicit paranoia that goes through a lot of those things; we've haven't studied it, but I assume it's high.

DR. IRVINE: The only reason I'm asking this is because as a psychometrician I found myself in a large cognitive department for many years. I've been doing what you've been doing; all of my cognitive friends have said "Have you read the materials on bias and reasoning because we publish most of it?" So my question follows. You have analyses that show that you have factorial strength in what you are testing. Have you thought of getting some of the other items or other kinds of bias and reasoning and putting them up against yours?

DR. JAMES: Yes.

DR. IRVINE: And have you done it?

DR. JAMES: No, these are all are opportunities. Another question has been, "Have you looked at it with projective techniques?" We've had one study where we've been able to do that. Right now, frankly this is not a magnet for funding. We just got a four million dollar grant to study climate, from NIMH. I haven't had an original idea in climate in 20 years. But it's a proven commodity so we can get a lot of money to do research. The Navy gave us a $100,000 grant, six or seven years ago, and that's about the only funding we've had for the conditional reasoning research.

DISCUSSION

DR. STRICKER: I wonder if the difference between the validity of your measure and the self- report measure in the meta-analyses is actually less than we might think. In the meta-analyses they are pooling across all kinds of different criteria and all kinds of variables within these different kinds of criteria, whereas you are focusing on a particular kind of variable.

DR. JAMES: Yes, it's not the fairest of comparisons. Personality ran into a lot of trouble in industrial psychology 25-30 years ago and in effect, they publicly called a moratorium on personality research. And the question one has to ask right now is, "Is personality in any better shape now than it was 25 or 30 years ago?" Well, there are probably some better measures and there's probably some better theory, but the validities are about the same. Before we get into another problem where we have to justify doing research on personality, which is where these meta-analyses are taking us, I think someone has to stand up and say something.

DR. MATTHEWS: The person actually asked the question I was going to ask, so to go back to the personality questionnaires and their validity. What the meta-analyses show, is if you take confirmatory studies or theory-driven studies, those give you validities of .4 or so (a paper by Joyce Hogan in *JAP* this year) which is about the same level as the implicit measure. So, the advance has not so much been in the personality questionnaires, but in getting better theory, which allows you to predict the criteria which the scale is going to predict.

DISCUSSION

DR. JAMES: I'm afraid I will have to disagree with you on that. I was very specific about saying uncorrected validities, and the validities that you are talking about are corrected validities. I have had a rather long history of debates with people concerning validity generalization and I particularly don't like the idea of correcting validities with an assumed reliability of a criterion of .52, which is what pumps those validities up. So, if you put those on a more even keel, either correct both, I would frankly be embarrassed to correct those, besides I won't use a .52 reliability. So if you look at the uncorrected and this differential, I looked at the Hogan and I think that is a little higher, but the uncorrected validity is still not that high. So you have to compare uncorrected to uncorrected or corrected to corrected, in that case the differential is there. Besides we also have the empirical data with four studies now, where consistently we find that difference in validity with the dominance studies.

DR. KYLLONEN: One more quick question … anybody?

DR. LEN WHITE: I wanted to ask and I think this is going to require you to speculate. Suppose you had instead choices of aggressive or non aggressive alternatives, based on the same justification mechanisms that varied moderately despite the specification, would you expect more of aggressive people, for them to be more able to more differentiate those levels of specification, than the less aggressive people?

DISCUSSION

DR. JAMES: Well, I do have some p-values on that new instrument and I'm framing that now in terms of the question you've asked. Off the of top of my head, no and when you look at the cross model in terms of proportions, no. But, I could be wrong on that, I would have to look at the data to answer the question. The discussion that we had yesterday is a real contribution and that was the idea that we could look at each of the alternatives, on a Likert scale and measure the probability that that is logical. What I like about that is it gives us the opportunity to add more alternatives to each scale, to enhance the measurement system. But no, I don't see those differences in specifications, but then again….

PARTICIPANT: It may make a difference?

DR. JAMES: Yeah, but it's set up so it's a qualitative difference, so the cognitive complexity would really be on one side of the aggressive people looking at various subtleties. But on the pro-social side, you have the same thing, all they have to do is reject, and it's set up to do that, then you could build those subtleties in on both sides.

<div style="border:1px solid black">

What is Beyond the Big Five?

Isms and Other Personality Constructs

Gerard Saucier

University of Oregon

Conference on "Emotional Intelligence:
Knowns and Unknowns" ETS, Princeton NJ

14 November 2003

</div>

WHAT IS BEYOND THE BIG FIVE? ISMS AND OTHER PERSONALITY CONSTRUCTS

DR. KYLLONEN: I want to introduce our next speaker, who is going to be taking us in a slightly different direction. By the way the theme of this particular session is things that are related constructs to emotional intelligence so we are touching on personality, and things that are related to personality. Next speaker is Gerard Saucier from the University of Oregon.

DR. SAUCIER: I'm aware that some of you may have been sitting in the same seat for two hours now, so if you want to take a 7th inning stretch, there won't be any patriotic songs. Pat said this was going in a different direction than the previous presentation; that is not entirely true, I noticed that the previous presentation involved beliefs and this particular presentation is also going to go in the direction of pointing the way in which peoples beliefs whether implicit or explicit can supplement personality and more traditional self report personality measures. I have become aware that an issue with regard to emotional intelligence is overlap with personality and that one critique that has been made particularly of self-report measures of emotional intelligence is that they may overlap with the Big 5, be highly predicted with the Big 5.

```
        The Study of Individual
               Differences

A Preliminary Basic Question:

Which individual differences are most worth
measuring?



•In personality

•In abilities

* And what about social attitudes, or
ideological beliefs?
```

What I want to do today is focus defining what the content of personality measures is as an area of individual differences and what sorts of aspects of personality or other important individual differences are not captured on current measures current popular measures. I must warn you that this is going to lead me into talking about people's attitudes about politics, religion and money and these are somewhat taboo topics and they have been that way in fact for personality psychologists but I think this presentation will indicate that we can avoid those topics somewhat at our peril.

For people in many research areas, attitudes of that nature are something of a nuisance, you would rather keep completely clear of that. But even if that is your attitude, it is useful to know how to measure them so you can control for them and make sure you have got some accurate measures. There are other aspects that psychologists should be interested in such as international affairs, violent social movements, relationship compatibility, where these kinds of contracts and measures may really be more predictive and useful than self report personality measures.

```
          The Study of Individual
                 Differences

A Preliminary Basic Question:

Which individual differences are most worth
measuring?



•In personality

•In abilities

* And what about social attitudes, or
ideological beliefs?
```

In the study of individual differences an important basic question is which individual differences are most worth measuring. I got into the study of individual differences by taking a class with Lou Goldberg about 15 years ago, and I recall the first day in class he started talking about something and this is exactly the question that I asked him, and it's been sort of prescient to what my research career has been about. It is an important question with regard to personality with regard to abilities and it's been asked a lot with regards to those areas. But I'm also going to get today into what individual differences are most important in regards to social attitudes or ideology beliefs. Now in determining which are the most important variables to measure, one recourse might be to authority which is to say you respect somebody as an expert so you believe what they about what the most important variables are.

```
      The Lexical Rationale for Variable
                   Selection:


  * The most important human attributes are
     encoded in language.

  * The degree of representation of an attribute
     in language has some correspondence with
     the general real-world social importance of
     the attribute

 AND…the most important attributes, or factors
     made up of attributes, should demonstrate
     cross-cultural generalizability
```

One alternative to this, is to study the language, because of the rationale could be that the most important human attribute tends to become encoded in language and the degree of representation of an attribute of language has some correspondence with it's social importance, therefore you could learn something about which is the most important variables to study by looking at what is emphasized in a language. This is a necessary by not sufficient criteria, it's not saying that all the important variables are a language, or you should limit yourself to what's in a language, it's more a matter of whatever is emphasized in language you know should be included in your model somewhere, supplemented perhaps probably by other things.

This is basically what's been called the lexical hypothesis, although I don't really think of it as a hypothesis, so much as a rationale for variables selection. An additional criterion would be or part of that hypothesis evolved is that the most important attributes or factors that you derive from many attributes should demonstrate some kind of a cross cultural generalizability.

```
    The Lexical Rationale for Variable
                  Selection:


  * The most important human attributes are
     encoded in language.

  * The degree of representation of an attribute
     in language has some correspondence with
     the general real-world social importance of
     the attribute

  AND…the most important attributes, or factors
     made up of attributes, should demonstrate
     cross-cultural generalizability
```

Now cross cultural generalizability, I consider an underused criterion and a lot of areas of psychology, we use a sort of coca cola model, which is to say that we develop our product in a western country and then export it world wide, with the assumption, what's valid for college students in Indiana is valid for everybody. The kind of approach I'm going to describe with regards to lexical studies of personality actually enables us to use a different approach, where we might be able to develop a model that works simultaneously in a wide variety of cultural settings. And we may be able to locate a model that's not primarily from one cultural area or one group of languages but rather it either is universal or fairly well distributed across global languages.

There have been by my count 14 lexical studies in different languages, these are studies where the personality terms typically adjectives are collected from a dictionary and they're reduced typically the most frequently used terms and then these terms are administered to participants to find out how people actually use them in every day descriptions of persons and then we take a look at which ones are redundant with each other and generate factors that can correlate with each other and get factors that way, of these studies, I said 14 of them, 10 of them have been studies of languages of European origin, which does have a bit of ethnocentric bias. I am engaged in some project to try to correct for that by studying particular African languages, to bring a kind of more global diversity to this.

```
    What We Learn From Lexical Studies of
       'Indigenous' Personality Factors
      (factor analyses of very large #s of descriptors)


  o "Big One" solutions yield an Evaluation
  (desirable versus undesirable traits) factor

  o "Big Two" solutions yield Social Propriety
  (socialization) and Dynamism factors (the
  saint, the hero)

  o Five-factor solutions most often yield the
  Big Five, but not always

  o Seven-factor solutions quite often
  (Filipino, Hebrew, and others) show an
  alternative "Multi-Language Seven" pattern…
```

But what I want to do here is summarize fairly quickly some extensive reviews of this literature that I've done with Lou Goldberg, these reviews have been published in a succession of books and journals. What you get really depends on how many factors you extract. If you just extract one factor which of course would be unrotated the universal pattern in all these studies, most of them have reported this is that kind of a Big 1 solution would be an evaluation factor, contracting desirable with undesirable attributes. That's also incidentally the Big 1 factor in Osgood's studies of judgments about the attributes about objects in general so there is probably a larger cognitive architecture going here.

A Big 2 solution which would be two rotated factors, this is also never been disconfirmed in any studies and most of them have reported something like this. That one of the factors will be social propriety, another could be labeled dynamism, social propriety would include things like, being polite and responsible and humble and patient, where as dynamism is more about being stimulating and interesting whereas social propriety is about morality, dynamism is essentially independent of morality.

```
      What We Learn From Lexical Studies of
         'Indigenous' Personality Factors

       (factor analyses of very large #s of descriptors)


o "Big One" solutions yield an Evaluation
(desirable versus undesirable traits) factor

o "Big Two" solutions yield Social Propriety
(socialization) and Dynamism factors (the
saint, the hero)

o Five-factor solutions most often yield the
Big Five, but not always

o Seven-factor solutions quite often
(Filipino, Hebrew, and others) show an
alternative "Multi-Language Seven" pattern…
```

Three and four factor solutions vary a great deal, actually even five factor solutions vary a good deal, but you do often find a Big 5 pattern and I'm going to assume with this audience that you know what the Big 5 is, but I will review what those factors are in just moment, the Big 5 has been found in these lexical studies, more often in languages of Northern European origin. I do want to point out if you take a Big 5 measure and impose it on some other culture by translating the measure, you can get the factor structure out of the other language.

In lexical studies, we studied each language on its own and looked for the indigenous structure, this is actually a more stringent cross cultural replicability criteria and the Big 5 did not fair as well using that more stringent criteria. Some recent studies are indicating that there are structures with more than five factors that are probably at least as replicable in these lexical studies as the Big 5, one that I've published is one that I called the multi-language 7, which I noticed from correspondence between Philippino and Hebrew lexical studies and that is in 2003 European journal of Personality article.

```
Multi-Language 7 - Big Five Factors


Gregariousness      |
                    |-> Extraversion
Self-Assurance |    |
                    |----> Emotional Stability
Even Temper     |   |
                    |-> Agreeableness
Concern for Others|


Conscientiousness -> Conscientiousness


Intellect        ----> Intellect


Negative Valence      ..............................
(Social
Unacceptability)
```
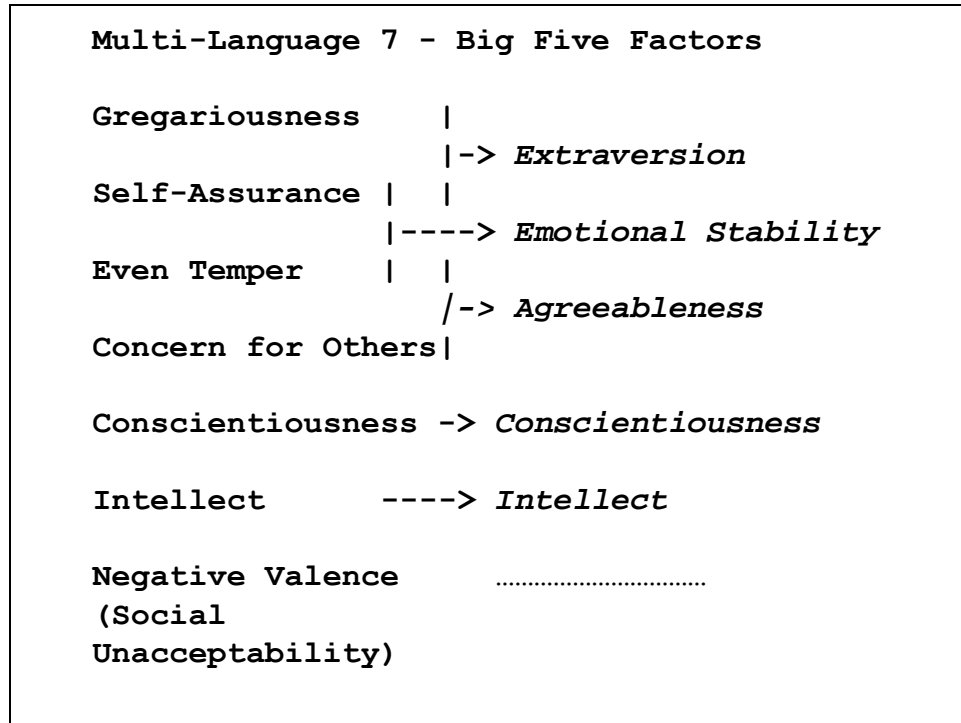
That particular structure resembles the Big 5 in some ways and is distinct in other ways; they both have consciousness and intellect factors. The multi-language seven has a negative valiance factor which has a number of socially unacceptable attributes correlating together with it, attributes like evil, good for nothing, retarded and corrupt and things like this.

So I have the Big 5 factors on the right and the multi language 7 on the left. It also divides these three kinds of affect interpersonal factors from the Big 5 somewhat differently into 4 factors. So that you get gregariousness and self assurance, even temper, concern for others, this is supposed to represent that the extraversion variance in this model is split into gregariousness and self assurance, emotional stability is split into self assurance and even temper which is kind of lacking a bad temper, so you have fear and anger actually separated in this model on the left. And agreeableness is split into even temper and concern for others. Another approach we can take in looking at content representation and of personality measures is to look at what source of personality variables might be well outside the Big 5.

What Traits (Represented in Adjectives) Are "Beyond the Big Five" (Saucier & Goldberg, 1998)?

…of 53 candidate outlier clusters, among the most reliable AND orthogonal were:

* Height (tall-short)
* Girth (chubby-skinny)
* Youthfulness (young-old)
* Religiousness (religious-nonreligious)

…**ideology or trait?**

With Goldberg, I did a study, published in 1998, where we looked at things that were really highly orthogonal to the Big 5. That devised 53 candidate clusters by studying a variety of adjectives, and the ones that were most, with the clusters that were not only orthogonal but also reliable, of course anything that was unreliable might also be orthogonal, but it wouldn't be of any importance. I have listed here the four that really was the best case for. One of them was height, which there are adjectives for, girth was independent of the Big 5, youthfulness and religiousness was virtually independent of the Big 5. Now this issue of religiousness as a trait started to raise some issues certainly for me about whether this was a personality variable or not, it is a matter of some controversy, whether this is ideology or personality.

```
     Personality and Ideological Beliefs
          (i.e., social attitudes)


  * Typically separated domains of variables

  * As expressed in adulthood, both are
     substantially heritable; beliefs show
     more shared family environment effects

  • Belief-related dispositions have powerful
     social effects (e.g. aggression/violence)
     and may be just as important

  * Not clear how much overlap between domains
     exists (in popular models and measures)
```

So in the next part of my talk, I'm actually going to go off in the direction to look at ideology variables and personality variables and how they compare to each other and how much they overlap. Traditionally social attitudes and personality have been separated. I'm not quite sure what to call this second domain, social attitudes, I don't like the title so much, ideological beliefs, I think it implies something, I don't think that's quite the optimal, if I had to devise the perfect label for it, it would be beliefs about what is true, or real, or good. But that's kind of long label, I'm still working on the best label for this domain. But I'm going to use these labels for it today, these have been typically been regarded as separable domains and one reason its long been assumed that the ideological belief or social attitude domain would not show inheritability would have strong environmental influence. But in fact the behavior genetic studies don't completely support that, they do show more of the tendency for belief to have shared environmental effects, at least in youth, not necessarily later in adult hood. But both are actually substantially heritable.

Of course a wide range of beliefs have not been studied yet. I'm going to use an extremely broad measure of ideologically beliefs later on here. I believe that most of those constructs I think heritability has not really been directly studied yet. This position that leads to beliefs are important variables that have powerful social effects. So I don't think they should be summarily disregarded. But it's not clear just how much overlap there is between these two domains, so I want to exert some things from a current paper of mine on this relationship.

Ideological Beliefs
(Social Attitudes)

Defining them: "Representations about what is true or real or good, held by an individual, though capable of guiding a group"

Measuring them: There is no comprehensive inventory, indeed often represented simplistically as a single continuum

---

First of all, I'm going to have to define what I mean by ideological beliefs or social attitudes and this is my definition. Representations about what is true or real or good held by an individual though capable of guiding a group, this definition is my own but it is based on a rather detailed survey of definitions in the psychological and sociological literature. In terms of the measurement of ideological beliefs I would say that the situation is somewhat more primitive that as the case for personality. There is no comprehensive inventory and indeed the domain is often assumed or thought to be represented by a single continuum, perhaps the left right, or kind of political and religious kind of continuum.

> ### Lexical Rationale Applied to Ideological Beliefs: Saucier's (2000) Study of English-language "-ism" Terms
>
> - Assumed the most important concepts in domain represented in words with —ism ending
> - Extracted 266 such terms from dictionary
> - Built items from content in the definitions
> - Constructed 400-item inventory
> - Administered to 500 university students (CA)
> - Found internally replicable 3 and 4 factor structures
> - In follow-up studies, the 4-factor structure has proved far more replicable

In order to kind of get beyond those kinds of assumptions, in a paper published in 2000 in *JPSP*, I reported on a study of isms terms in the English Language. That was just in one language, but the results have been well replicated in a yet so far unpublished study in Romania, done by Stephen Kraus. This was based on the idea that for the beliefs ideological belief domain that you could capture the most important variables by looking at words that end in I S M, ism.

So what I did with my research team is to go through the American Heritage Dictionary of the English Language, we extracted 266 isms terms that did refer to beliefs, now we did have to leave out some terms that end in ism such as metabolism or alcoholism, I don't want to short strip those things, I'm a strong believer in metabolism but it wasn't really a kind of belief that there would be individual differences in, so we were concentration on ones that were actually belief terms. The strategy here was not unlike the lexical studies of personality, we did not administer the actual ism term to two people, partly because some of them are a bit unfamiliar and also we expected a tendency where many people would say that I don't believe in any isms, I believe what is true they are not isms.

```
        Lexical Rationale Applied to Ideological
          Beliefs: Saucier's (2000) Study of
            English-language "-ism" Terms

  • Assumed the most important concepts in
    domain represented in words with —ism
    ending
  • Extracted 266 such terms from dictionary
  • Built items from content in the definitions
  • Constructed 400-item inventory
  • Administered to 500 university students
    (CA)
  • Found internally replicable 3 and 4 factor
    structures
  • In follow-up studies, the 4-factor
    structure has proved far more replicable
```

But what we did instead is that we looked at the dictionary definition and we turned the definition into questionnaire items. For example a hedonism item was definition included the idea that pleasure is the chief goal of life, or one definition was the belief that pleasure is the chief goal of life, we just made it into an item, pleasure is the chief goal of life; do you agree or disagree? By a direct a method as possible we just converted the definition into an item. Out of that we constructed a 400 item inventory, it's more than 266 because we had to, many of the definitions were multi, there was more than one definition for the term and some of the definitions were so long we had to split that we had to split them into multiple items then recombine them later on.

A good example of that is environmentalism, which has two very different meanings, environmentalism can mean you want to protect the environment from pollution, or it could mean that you believe that environment is more important than genetics. Those two are completely separable forms of environmentalism.

So we administered the 400 item inventory to 500 university students and found an internally replicable 4 factor structure, replicable across halves of the sample. And in follow up studies I compared various structures from the initial study and it was the 4 factor structure that proved the most replicable.

```
              Four Replicated Isms Factors

    Alpha (α) -- Religious Orthodoxy (encompassing
        fundamentalism versus secularism)… religiousness,
        correlated with conservatism and authoritarianism

    Beta (β) – Unmitigated Self-Interest – advocating
        unfettered exploitation of opportunities by self or
        close in-group… correlated with materialism,
        hedonism, ethnocentrism

    Gamma (γ) – Protection of Civil Order (status quo
        polity, common civil institutions)

    Delta (δ) – Unorthodox Spirituality and Mysticism –
        emphasizes subjective experiences, including
        paranormal ones
```

So let me just review what these factors are so you know what the content of some the later analysis are going to be dealing with. My initial labels for these were cautious, and I used Greek letters because I didn't want to commit myself too early to the meaning to labeling and this was partly informed by the experience with the Big 5 where I think some of the labels are not so great. But I would characterize the first biggest factor as religious orthodox, which I call the alpha factor. It does involve religiousness and it encompasses a stark contrast between fundamentalism on one pole and secularism on the other and is correlated as I demonstrated in the 2000 study with measures with conservatism and authoritarianism actually quite very substantially. In fact actually if we go back to this factor, this is the factor that has been most emphasized in previous measures of social attitudes. So that if you had put all of them together you would probably get one big factor which was this and you wouldn't get so much as the next three factors that I'm going to subscribe, one of which is I call Beta, it could be labeled as unmitigated self interest, a belief that you should pursue your self interest sort of no matter what versus belief in the limit on self interest, so through the core of logic through the high Beta person would be that they advocate unfettered exploitation in environmental opportunities by their self or by self close end group such as their nation or their family against others and some of the correlates are actually components of this factor are materialism, hedonism, and ethnocentrism, from an evolutionary standpoint, this is probably maximizing short term fitness but disregarding considerations of long term fitness.

```
        Four Replicated Isms Factors

  Alpha (α) -- Religious Orthodoxy (encompassing
     fundamentalism versus secularism)… religiousness,
     correlated with conservatism and authoritarianism

  Beta (β) – Unmitigated Self-Interest – advocating
     unfettered exploitation of opportunities by self or
     close in-group… correlated with materialism,
     hedonism, ethnocentrism

  Gamma (γ) – Protection of Civil Order (status quo
     polity, common civil institutions)

  Delta (δ) – Unorthodox Spirituality and Mysticism –
     emphasizes subjective experiences, including
     paranormal ones
```

A third factor labeled gamma really refers to beliefs in the civil rather than a religious kind of order, almost a civil religion to use the term of Russo, so protection of civil order is the emphasis here, and that is the status order quo political orientation and a common civil institutions like the government and the constitution and so on. And the fourth factor is Delta which is unorthodox, spirituality and mysticism as contrasted with the more orthodox religiousness of the Alpha factor, the emphasis here is one subjective experiences, rather than kind of an extrinsic approach to religion and it includes a tendency to believe in paranormal experiences as well.

```
          Derivation of public-domain Survey
            of Dictionary-based Isms (SDI)

   •  Delineates lower-level constructs, not just
      broad factors, each an item composite with
      ~4 items, keying in both directions
   •  Each based on at least two of the original
      400 isms items
   •  Created new items to allow balanced keying
   •  Administered to new college-student sample
      (N=383, OR)
   •  Finalized using data from third, community
      sample (N=703, OR)
   •  In end: 48 item-composites, 12 assigned to
      each factor, α .51 to .91 (half above .73)
   •  For each factor, primary (p; more univocal)
      and secondary (s; less univocal) subscales
```

In order to develop better measurement in this domain, I wanted to develop a measure that had a lot of fidelity with regard to lower level constructs, not just broad factors. So I constructed a large number of item composites that have roughly four items, and I found it very important to make the direction of keying balanced within these item composites because attitude measures have a rather troubled history of being affected by acquiescence individual differences in acquiescence, and I wanted to get a control on that. Each of the item composites are based on at least of the two original 400 ism items, but I did generate additional items in order to allow for balance keying, that was necessary. So I call this the survey of dictionary based isms, it is a public domain inventory, not proprietary. I administered the initial draft of this to new college students sample and finalized it with a third community sample. Has 48 item composites, 12 assigned to each factor with Alphas between .51 and .91 and I do divide these subscales into primary and secondary subscale based on their tendency to be always loading less highly on the factor versus shifting around from one factor to another.

**Empirical Study: Relation of
Personality and Ideological Beliefs**

*2000 study indicated isms quite independent of
    adjectival Big Five -- but what about other
    personality models?*

METHOD

**N=703, Eugene-Springfield community sample, 57%
    female, average age 59 (range 20s to 90s)**

**Administered 'Survey of Beliefs and Opinions'
    including the SDI items**

**Had been administered numerous personality
    inventories in previous eight years**

I have a current paper that I'm finishing off about the relation between personality and ideology beliefs in the 2000 study I found that the isms were quite independent of the adjective Big 5, that was just one brief measure that I used in that study. I am going to describe here a much more involved study looking at a large variety of personality measures, using the Eugene Springfield community sample. And I administered to this sample a survey of belief and opinions which included all of the SDI items that make up those 48 clusters. This sample had been administered numerous personality inventories in the previous 8 years so it enabled some useful comparisons.

```
 Personality Inventories Administered


 * 10-adjective measures for each factor:
      Big One (Evaluation)
      Big Two Dynamism
      Big Two Social Propriety

 * 100 Big Five marker adjectives (of
 Goldberg)

 * 60-adjective marker set for the Multi-
 Language Seven
```

I didn't look just at the big five with regards to adjectives; I also used 10 adjective measures that I used in another published study for the Big 1 and for the Big 2 factors that are lexical. I used Goldberg's hundred Big 5 marker adjectives for the Big 5, and I used the 60 adjective marker set for the multi language 7.

```
Personality Inventories(and # of scales)


 California Psychological Inventory (CPI; 3)

 NEO Personality Inventory (NEO-PI-R; 5)

 Six-Factor Personality Questionnaire (6FPQ)

 Hogan Personality Inventory (HPI; 7)

 Temperament and Character Inventory (TCI; 7)

 Multidimensional Personality Questionnaire
    (MPQ; 11)

 16 Personality Factors (16PF) questionnaire
```

In addition, I looked at 7 different well known personality inventories, and I used broad at least fairly independent factors that can be derived from these measures. So that Harrison Gough CPI I used the 3 Factor scales, for the NEO inventory of I used of course the 5 domain scales, Doug Jackson 6 factor personality questionnaire I used these 6 broad scales, Hogan personality inventory, I used 7 broad scales, Cloninger's temperament and character inventory I used the 7 broad scales, for Tellegen's MPQ, I used the 11 Broad scales, these are getting to be a little bit correlated, but I kept them, kept all 11 of those, and with the 16 PF I used all 16 partly because that is the most commonly used scales, although there are broad scales for the 16 PF, to measure the isms factors I used the 48 SDI item composites and generated factor scores from them.

```
        Isms Factors from 48 SDI Item-
                   composites


   •  Factor scores from four-factor varimax-
      rotated solution (principal-factors
      extraction method)

   •  Eigenvalues consistent with optimal 4-
      factor solution

   •  All primary (p) subscales loaded on
      expected factor

   •  Reasonable representation of 4 isms factors
```

The Eigenvalues from that factor analysis were consistent with a 4 factor solution being optimal. And all of my P subscales or the primary subscales loaded on the expected factor. And I do think this is a reasonable representation of the 4 isms factors.

```
      Single Bipolar Continuum of
          Ideological Beliefs:
      Republican versus Democratic

Items:
  "Politically I
  favor the
  Republican party"
  and "Politically
  I favor the
  Democratic party"
• Second response
  subtracted from
  first
• An α of .93
```

Now I mentioned before that there are often ideological beliefs are assumed to be represented by a singular bipolar continuum which is in America would correspond to Republican and Democratic Party. So I did have two items on my inventory that refer to this: "Politically I favor the Republican Party" and another one for the Democratic Party. I simply had a simple measure of this by subtracting response of the second from the first, this is all using a 1 to 5 multi point rating scale and it turns out they are very highly correlated around negative .8 I believe, and the Alpha was .93 for only a 2 item scale, pretty reasonable way of measuring this.

```
                    Analysis Strategy

   * Each ideological belief variable
       - American Political Party Affiliation
       - 4 isms factors
     served as criterion variable in multiple
     regression

   * The set of scales on each inventory served
     as a set of predictors in each regression
     procedure

   * Adjusted (shrunken) R values examined –
     corrects for different number of
     predictors and sample-size variations
```

So I am going to be comparing that particular single bipolar approach to ideological beliefs with the isms factors here, so my analytic strategy here was to use multiple regression for each of the ideological belief variables that would be the Republican minus Democratic measure and each of the four isms factor was a criterion variable in multiple regression. And the predictors for each regression were all the scales from one of the personality inventories, for example all of the MVQ scales would be the predictors and one of the ideological belief dimensions here would be the criteria, I'm going to look at how big the multiple R's are, but I am going to use adjusted or shrunken R values because I've got different numbers of predictors with these different inventories and also because I'm using deletion procedures I'm getting some sample size variations and I wanted to make everything more comparable.

Adjusted R for Lexical Personality Models Predicting Ideological Dimensions

Ideological Dimension	Big1	Big2	Big5	ML7
Republican – Democratic	.00	.08	.28	.34
Alpha (Relig. Orthodoxy)	.00	.04	.25	.40
Beta (Self-interest)	.14	.21	.27	.23
Gamma (Prot. Civil Order)	.23	.15	.33	.41
Delta (Unorth. Spirit.)	.06	.12	.31	.31

Here are the adjusted R values that I get in this analysis for using the person the lexically derived personality measure, these are adjective measures in predicting the ideological dimensions. You can see that the Big 1 in some cases has a zero correlation, it doesn't predict some of the ideological dimensions at all. Generally as you add more factors to the lexical model, you get increased prediction of the ideological dimensions. But none of these coefficients is the highest is .41.

Adjusted R for Questionnaire Models
Predicting Ideological Dimensions

Ideol.Dim.	CPI	NEO	6FPQ	HPI	TCI	MPQ	16PF
Rep.-Dem.	.26	.39	.24	.20	.25	**.46**	.39
Alpha	.36	**.47**	.32	.33	**.63**	**.65**	**.55**
Beta	.33	.33	.20	.20	**.50**	.30	**.53**
Gamma	.38	.36	.33	.38	.38	**.46**	**.48**
Delta	.09	.44	.19	.29	.44	**.48**	.31

If I then turn to the questionnaire models I start finding many of the multiple R's are now exceeding .41, in fact about a third of them are here, so we do get higher multiple R values even though we are using shrunken R values, with the questionnaire model, so that is true for some questionnaires and not for others, the CPI and the 6FPQ and the HPI really don't generate higher values than the lexical models do, but the NEO, the TCI, MBQ and 16PF all generate higher multiple R values. That indicates that they have more overlap with ideological beliefs than the other measures do. But something's going on here that's making these measures really encroach into the ideological belief domain, more than other personality measures. And it's more true if you notice here for the Alpha factor. That's true, the highest always for the Alpha factor for all of those four inventories that I targeted, however there are some substantial multiple R's for the other ideological dimensions as well. I was somewhat surprised to find that you could predict voting preferences probably to some degree using personality measures.

```
Summary of Multiple Regression Results

Coefficients in table are higher for
questionnaires, especially 16PF, MPQ,
TCI, and NEO-PI-R

Tendency — though not monotonic —
more scales in inventory, higher R for
ideological belief dimension

Religious orthodoxy especially well-
represented in these inventories
     MPQ Traditionalism
     TCI Self-Transcendence
     16PF Rule-consciousness
     NEO-PI-R Openness (low)
```

So I think I've summarized this already that the coefficients are higher for questionnaires especially for the 16PF, MPQ, TCI, and NEO, in the case of the NEO it's really the openness to experience scale, which is generating this, because it has the same pattern of correlation with the isms factors as that republican democratic measure did. That reflects the fact that high openness to experience tends to predict to almost R (square) = .4 that one will vote Democratic, as opposed to Republican. Religious orthodoxy is especially well represented in these inventories but to a lesser degree but we do find other dimensions represented.

```
    Correlations of Specific Personality
         Scales with Isms Factors

 •  Most correlations low, but highest are:


    ALPHA (Relig. Orthodoxy)
       .63 MPQ Traditionalism
       .60 TCI Self-Transcendence
       .45 16PF Rule-consciousness

    BETA (Unmitigated Self-Interest)
       -.41 TCI Cooperation

    GAMMA (Protection of Civil Order)
       .39 CPI Norm-favoring

    DELTA (Unorthodox Spirituality)
       .44 TCI Self-Transcendence
       .42 MPQ Absorption
```

If you look at the specific personality scales that are generating these correlations we see that MPQ has a traditionalism scale TCI has a self transcendent scale, 16PF has a real consciousness scale those are really kicking those multiple R's up quite a bit. NEO openness is actually correlated just below these levels with all of the factors but especially with alpha. You also see cooperation, Norm favoring and absorption scales being correlated with the isms factors.

```
        Personality and Ideology:   Are They
                One Thing, or Two?

• Overlap is modest for the two-party continuum

• Higher when four isms dimensions considered

• Especially high for Religious Orthodoxy (α)

• Overlap modest for lexical-based scales

• But higher for questionnaire models that
include scales like Traditionalism, Self-
Transcendence, Openness, etc.

• To a degree, these domains already intertwined
```

So I can ask the question of whether personality and ideology are one thing or two and you would have to say they tend to be two things more if you think about the two party continuum it is somewhat higher if you take into account all four isms dimensions, there is more overlap with religious orthodoxy, so you get that kind of content encroaching into personality measures more than other kinds of contents, lexically based scales have rather less overlap and so to a degree these domains are already intertwined and you use a lot of the personality measures you're already involuntarily perhaps measuring ideological beliefs to a degree.

```
     Do Isms Factors Offer Incremental
     Validity Over Personality Factors?


•Isms factors largely independent of lexical
personality factors

• But this does not establish their importance –
their validity

• Previous literature (correlated constructs)
already establishes good case for Alpha (Relig.
Orthodoxy), less so for other factors

• An informative criterion test would be with
measures relevant to "emotional intelligence"
```

Now there is a question of whether these ideological beliefs are a nuisance variables or whether they actually add something interesting to prediction of behavior. This is what I want to get to next, and this will bring us more directly into emotional intelligence topic. Largely the isms factors are independent of the lexicon based personality factors. But just because they are independent of them doesn't mean that they're somewhat useful beyond those factors. There is already a pretty good case, that for constructs associated with alpha associated with Alpha such as authoritarianism or conservatism that those are useful variables for some things, we don't much about the validity of other factors. So what I want to use here as the last part of my talk is a criterion test where I use self report measures that are supposed to be relevant to Emotional Intelligence and I am going to hit personality measures against these ideological belief measures and see what if the beliefs actually add to the predictions over personality measures.

```
          Criterion Self-Report Measures


   Toronto Alexithymia Scale (20 items)
      …used three subscales…
      Externally-Oriented Thinking
      Difficulty Identifying Feelings
      Difficulty Describing Feelings

   Barchard's Emotional Intelligence scales
      Attending to One's Own Emotions
      Negative Expressivity
      Positive Expressivity
      Responsive Joy
      Responsive Distress
      Empathic Concern
      Emotion-based Decisionmaking
```

The Emotional Intelligence relevant self report measures fall into two groups, one is the Toronto Alexithymia Scale which I'm using the 20 item version, and I'm going to use not the overall score, but the three subscales scores, externally oriented thinking, difficulty identifying feelings and difficulty describing feelings, also Kim Barchard in her Doctoral Dissertation devised some Emotional Intelligence self report scales. I don't want to represent that she thinks that this is the best measure of Emotional Intelligence, actually she concluded Cognitive Measures that treated as ability would be superior, these she thought had too much overlap with personality. But these are the scales that she used, Tending to Ones Owns Emotion, Negative Expressivities, Positive Expressivities, Responsive Joy, and Distress and Empathic Concern and Emotion Based Decision Making. Some of these are related to some of the emotional skills that Paul Ekman talked about, such as being able to identify feelings, and paying attention to your emotions, although this isn't self report format these are conceivably relevant to emotional skills.

```
   Method for Incremental Validity Study

• 522 community-sample participants completed
all measures

• Lexical personality scales added as blocks
with increasing numbers of factors, to examine
their relative incremental validity:
Big One -> Big Two -> Big Five -> ML7

• Final block: the four isms factors

• Examined change in R-squared values for each
new block of variables, for each of 10 EI-
relevant measures
```

My method here was to use 522 participants from that community sample and to actually in addition to comparing the attitudes with the personality skills I am going to look at how adding more lexical personality factors add to the prediction to what degree it adds to the prediction to emotional intelligence, so I'm really comparing Big 1, Big 2, Big 5, and multi language 7 representations of predicting self report emotional intelligence measures. And the final block is going to be the four isms factors, so I'm going to be capturing the change in our square that you get from adding the isms factors after the personality variance has already been taken out. I am going to be looking here at change in our square values for each of the 10 EI relevant scales separately, these are the R squared change values I get and the * here indicated it was significantly change in our R-square.

```
   Hierarchical Regression Results

                    Big1 Big2 Big5 ML7   Isms
TAS
Ext.Oriented       .01* .07* .12* .04*   .11*
Diff.Id.Flgs.      .07* .03* .13* .04*   .03*
Diff.Des.Flgs.     .06* .16* .05* .04*   .02*


Barchard scales
Emot.Decisions     .00  .03* .14* .05*   .13*
Empathic Concern   .05* .06* .10* .04*   .10*
Attending Emots.   .00  .02* .17* .03*   .06*
Responsive Joy     .10* .05* .12* .08*   .04*
Respon. Distress   .00  .02* .22* .07*   .04*
Positive Exprvty   .08* .11* .12* .06*   .04*
Negative Exprvty   .00  .13* .11* .05*   .02*
```

And it turns out that adding more personality factors in essentially every case adds significantly to the prediction of these EI measures; adding the isms factors also in every case added significantly to their prediction. You can get more prediction by adding more personality factors but also the isms were adding, but I do want to point out that the change in R squared is especially large for three of the EI measures, but it is relatively low for some of the others. The externally oriented thinking isms actually added quite a bit to the prediction of that, also emotional decision making and empathic concern. Overall the personality variables the multiple R was .53 and adding the isms factors raised this value to .58. If you turn the analysis around enter the isms factor first they generate a average multiple R across these scales of .37 which is lower but for those three highest coefficients of externally oriented emotional decision making and empathic concern, actually isms on their own predicted the EI scales more highly than the personality measures did.

Results of Hierarchical Regressions

Typically…

~11% of EI-scale variance accounted for with only one or two-factor personality models

Adding Big Five more than doubled the R^2

Adding Multi-Language 7: Further improvements

Adding isms factors __always__ significantly increased R, with 2 to 13% change in R^2

This increase highest for EI scales involving attention, preoccupation, decision processes

You can get about 11% of the EI scales variants with the 1 or 2 factor personality models adding the Big 5 more than doubled the R squared, it increased beyond 22% and the multi language 7 would add significant further improvements, the isms factor always significantly increased the R but how much varied from one scale to another. And the ones where it did have more of an effect related to attention preoccupation and decision making which may be more cognitive aspects of EI.

```
          Isms Factors and EI Scales

   • Beta (unmitigated self-interest) had
     significant negative association with
     all ten EI scales

   • Delta (unorthodox spirituality) had sig.
     positive association with 7/10

   • Gamma (protection of civil order) sig.
     associated with 7/10

   • Alpha (religious orthodoxy) predicted
     none…Alpha does predict other things!
```

I took a look at which of the isms factors was actually contributing to this improvement and prediction of the EI scales. And it was never the Alpha Factor that was doing so. It was in fact the Beta factor was doing it in every single case, all ten scales, it had a significant Beta coefficient. The Delta in 7 out of 10, the Gamma in 7 out of 10, and the way it goes is that high Beta belief and unmitigated self interest tended to be correlated with lower scores on the emotional intelligence measures.

```
      Two SDI Scales Frequently Adding to
            Prediction of EI Scales


  Delta p1:   Intuitionism


  Example item:   "Enlightenment can be gained
  through meditation, self-contemplation, and
  intuition"


  Beta s1:   Physicalism


  Example (reverse-keyed) item: "Physical laws
  cannot explain some mental phenomena"


  …Both represent positive (versus indifferent)
  valuation of internally-oriented thinking
```

I wanted to go to even a more detail so this the result of step-wise regression where I looked at which of the item composites were actually most contributing to the prediction of the EI measures, and these are the two that stood out, so when there was a high increment of prediction these were always in the equation, pretty strongly. One of them was the intuitionism scale an example item, a belief that you can gain enlightenment through meditation, self contemplation and tuition. The other one is physicalism scale which would be choosing a reverse keyed item, because the reverse keyed item that would be the one that beta 1 would be positively correlated with the EI scales, belief that physical laws cannot explain some mental phenomena this actually helpful at least to me in understanding what is going on here because both of these represent a valuing of internal oriented thinking as opposed to only focusing on external. It may be that emotional intelligence; one contributory or precondition is the belief that subjective experience and what's going on internally is important. And that is not going to be captured on personality measures.

Important Caveat!

The "emotional intelligence relevant" scales employed here were all self-report, and not maximum performance tests.

One caveat is that these are just self report measures not maximum performance tasks so they subject to all the limitations to such measures.

Overall Conclusions

- **The Big Five doesn't provide a comprehensive account of individual differences**

- **Lexicon-based personality dimensions (like the Big Five) are generally independent of ideological beliefs**

- **Questionnaire models of personality often contain certain scales rather saturated with ideological-belief variation**

- **Ideological-belief dimensions (represented by isms factors), added to personality measures, increase the prediction of important criteria (including some related to EI)**

The Big 5 does not provide a comprehensive account of individual differences, I think I've given some demonstration of that although Lexicon based personality dimensions are generally independent of ideology beliefs the questionnaire models are not always so and the ideological belief dimensions when you add them to personality measures increase the prediction of important criteria, including some that are related to emotional intelligence so one of the contributions that this may make to the literature on emotional intelligence is just indicating that you can't disregard the importance of beliefs. Peoples beliefs systems in terms of how they handle their emotions.

Acknowledgements:

Thanks for great research assistance to:

Tarik Bel-Bahar
Teresa DaVigo
Amber Keele
Ching-Ting Shen
Sasha Smith

<div style="text-align: center;">

DISCUSSION

</div>

DISCUSSION

DR. KYLLONEN: We have time for some questions.

DR. SCHERER: One of the things that I was particularly intrigued by is the possibility that the isms factors that you identified might play a role in some of the biases and appraisals that I tried to describe yesterday, particularly paranormal beliefs in terms of, for example, attribution of responsibility, particularly in terms of causal responsibility. Do you think that there is any evidence that you are aware of, or do you have any hypotheses, concerning the effect on such appraisal factors?

DR. SAUCIER: No, I am not aware of such evidence, I think there has been a dearth of studies in this area; ideological beliefs have been separated off into a domain for political psychologists and sociologists maybe, so we haven't really looked at that too much. Just from the content of the measures you would expect a good bit of that.

DR. SCHERER: We did a little study where we tried to measure paranormal beliefs in 3 factors, one was divination, and people believe in divination practices, they actually think that there is less chance in appraisal. A little bit of evidence that it might work, could be very interesting.

DISCUSSION

DR. JOHNSON: What words happen to have isms in them and to what extent are they a happenstance of the English Language? What likelihood do you think there is that you are overlooking important belief systems by focusing on these words?

DR. SAUCIER: There is a good deal of likelihood that I'm overlooking the belief system, but I am correcting for that, because what I've done is I did a survey of the literature and I have 42 additional item composites that were actually administered to the same sample. I just haven't analyzed that data as much, the full comprehensive measurement that I would have right now would have ninety item composites, and that includes things like social dominance orientation which is not terribly well captured there. I have an authoritarism scale, I have some things related to attribution, why poor people are poor, lots of other variables like this. I am making an attempt to come up with something that fills in the holes.

DR. KYLLONEN: Another question. The question is there are a lot of social attitude surveys out there and maybe when you responded to Dr Johnson's question you were saying that is what you were doing in the supplementary measures. I guess my question is how much do the isms factors capture what the existing social attitude scales survey, and these dimensions already say?

DISCUSSION

DR. SAUCIER: To get a really get a good answer for that, I would have to administer all those social attitude measures, which I haven't really had the resources to do. What I did is this: I looked at those measures and figured out what core items were and more or less, adopted items like that into these item composites. But I haven't looked at that directly. I can say that if you factor analyze all 90 item composites -- the isms ones plus the additional ones -- you can get a couple of additional factors that are fairly independent of the first four, one of which has to do with belief whether big government is good or helpful, and another has to do with whether you want to treat outsiders in society harshly or kindly. So you do get some other kinds of content that way.

DR. KYLLONEN: One other question. You mentioned the first factor was very important relative to the others, indeed you suggested it was larger. How much larger?

DR. SAUCIER: It's not hugely larger, it's every factor analysis I've done, it's always been the first one. But it's not twice as large, even looking at the unrotated Eigenvalue; it is never more than twice as large as the second.

DISCUSSION

DR. SIGEL: You mentioned that these belief systems are what you are getting inside ideology. How do they relate to any social behavior, for example, voting behavior? Do you find any relationship between them?

DR. SAUCIER: Social behavior for example which?

DR. SIGEL: Voting behavior or other kinds of thing?

DR. SAUCIER: Voting behavior, I haven't personally looked at that relationship now, this Republican-Democratic continuum that I came up with should predict voting behavior pretty well unless I'm getting a lot of lying. The isms factors are correlated, you can a multiple R, I believe around .7 for predicting that particular continuum from the isms factors. Now this has raised one issue for me, which I've considered that it might be useful to have a model at least for American purposes, where you have one factor, be that political continuum and the other factors are completely independent of that. And it is clear that you can get three factors, which are completely orthogonal to that political continuum. In that model, one would predict voting behavior very highly and the others would maybe predict specific issues but not so much others.

DISCUSSION

DR. SIGEL: A related questions in terms of your notion of ideology and ideological beliefs. Is your view on belief a kind of a singular, univariate thing or is it componential? How would you define belief in a componential way, because there are other factors that are related to it?

DR. SAUCIER: How would I would I find beliefs in such a way?

DR. SIGEL: Components? Let me explain. For example, in the parent child literature there is a lot of stuff dealing with beliefs very differently than what you're dealing with. For example, a belief is more than just what you're saying. A belief involves also affect, knowledge and so forth, all of which are tied into this belief, but you're using belief primarily as the definition you presented, right?

DR. SAUCIER: It emphasizes rather broad global beliefs, about what is true, real, and good. You can devise beliefs about very specific kinds of entities, kind of endlessly and I can't guarantee that these are going to comprehend those, they probably don't completely. For example, the previous presenter had some, had conditional reasoning items, and I was able to make out for most of his implicit beliefs of aggressive people, I could come up with explicit belief items that I could put on inventory like this. So its' method that is fairly adaptable but it is not comprehensive of all beliefs. I agree! I tend to focus on broad, highly generalizable types of beliefs.

DR. KYLLONEN: Okay, thank you.

Social/emotional intelligence:
From factors to functions

Maureen O'Sullivan
University of San Francisco

SOCIAL/EMOTIONAL INTELLIGENCE: FROM FACTORS TO FUNCTIONS

DR. KYLLONEN: I'd like to now introduce Maureen O'Sullivan, who is going to discuss "Social-emotional intelligence: From factors to functions."

DR. O'SULLIVAN: A rose by any other name would smell as sweet, but if we call emotional intelligence by another name, would it sound as sweet? If we called it emotional consensus, or emotional interest, or agreeableness or emotional competence, as Klaus suggested yesterday, would it still have that zing, which has all of us here this weekend.
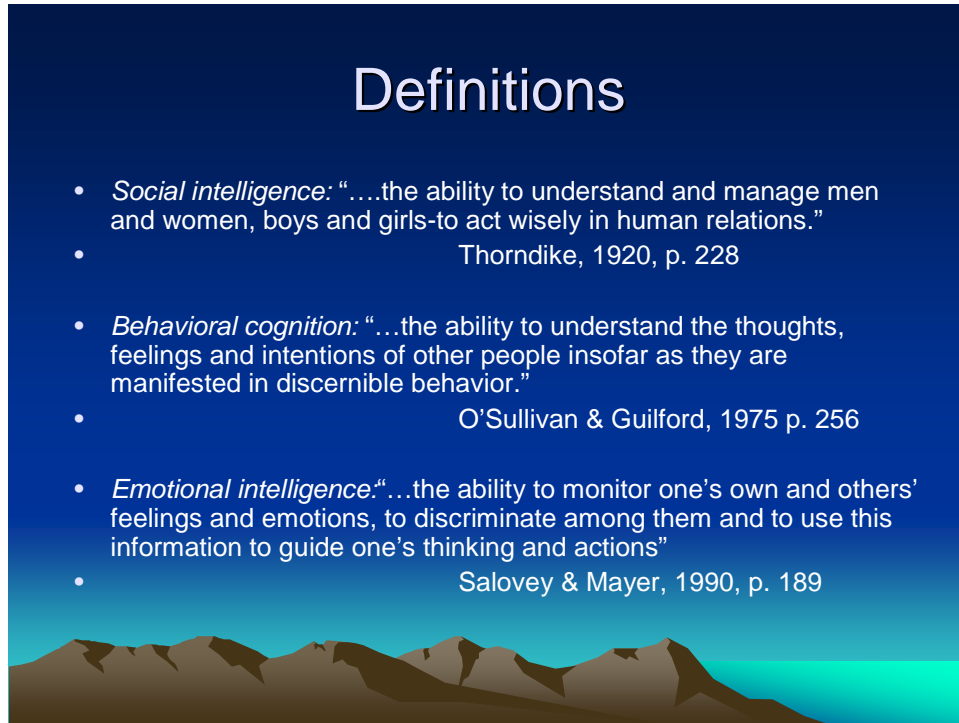
Psychologists have been interested in this general area of social, emotional, behavioral intelligence for a very long time. For example, I have up here some definitions, dating back to 1920, which is the earliest I could find for psychologists interested in this general area.

E. L. Thorndike, said that there were three different kinds of intelligence, one of which was social, he disagreed with Klaus Scherer, and thought there was such an intelligence.

In his 1957 structure of intellect model, J. P. Guilford suggested that 30 of his 120 factors of intelligence were related to what he called behavioral intelligence, the definition on the screen relates to our study of behavioral cognition or understanding other people, just a small part of that domain. He had hypothesized 24 other intellectual abilities, including behavioral creativity, behavioral appropriateness and so forth.

The last definition is Salovey and Mayer's original definition of emotional intelligence, now certainly there are differences among these definitions and the most striking one and perhaps the most important being the inclusion of self knowledge in the Salovey-Mayer definition.

But what is so curious is that if we look at how these definitions and how they were measured, we would find very little difference among them. Thorndike's 1920 article is illustrated with photographs of a woman posing facial expression of emotion. Many of the O'Sullivan and Guilford test assess the ability to recognize facial expressions as does many current emotional intelligence tests. So although the construct is differently named in terms of measures there is considerable overlap in terms of how it is that they are assessed.

So I find it really curious that we have all these different names and yet we have the same way of measuring them. Now certainly there are differences. Some use expert opinion, some use consensus, but essentially they are all looking at facial expressions of emotions and recognizing them. So different from what Bill Revelle talked about yesterday, where with psychometricians you have the same name, extroversion, and yet you have very many different ways, and everybody knows even though you're using the same word you're meaning different things. Yesterday when Drew Gitomer opened this conference he noted that early attempts to measure social emotional intelligence focused on abstract, consideration of factorial uniqueness; early researchers attempted to show that there was such a thing as social intelligence and it is still a question for some people whether this domain really involves an intelligence, a skill or an ability area.

So, early researchers were interested in showing factorial separation from measures of other intellectual abilities and after many years of trying they did so. In 1947, Wedeck, a student of Burt, demonstrated the existence of three factors of what he called psychological ability. These factors were orthogonal to two other intelligence factors. In 1975 O'Sullivan and Guilford demonstrated.six behavioral intelligence factors. The1970's also saw the publication of several tests of the ability to, as they called it then, decode non verbal behavior. But these tests were about recognizing facial expressions of emotion, more broadly defined.

These tests of non verbal behavior included Ekman and Friesen's affect recognition test (BART), a precursor of the METT that Paul Ekman talked about earlier. Rosenthal's Profile of Non verbal Sensitivity (the PONS) and Archer and Akert's Social Interpretation Task. None of these tests was significantly correlated with verbal or general intelligence.

So one of our knowns in this area is that it is possible to come up with something that relates to social or emotional intelligence, that does not measure g. Recent reports of positive correlations aomng modern emotional intelligence tests still left us in the dark about whether the shared variance between them is due to sensitivity to non verbal behavior, awareness of what the appropriate situation is in which to show such behavior, consensual similarities with other test takers, or some combination of the three.

As long ago as 1955, Cronbach called for the discrimination of what he called differential accuracy, differentiating one person from another in terms of understanding them, and stereotypic accuracy, understanding commonalities among a group of people. The clarification of this distinction in many current emotional intelligence tests, I believe, is one of our unknowns.

In the 1980's, Sternberg introduced what he called his triarchic theory of successful intelligence. One of the most interesting aspects of this theory is its underlining the importance of understanding situations and contexts. Sternberg and his colleagues developed measures of what he called tacit knowledge; one of our colleagues called it common sense yesterday, the ability to understand the requirements for success in a particular environments such as the military, the business world or the academic world. Data I will present later are consistent with this theory, so I think that we can safely say that one of our knowns is that there such a thing as tacit knowledge, that people differ in the ability to understand and to master to the requirements of different situations.
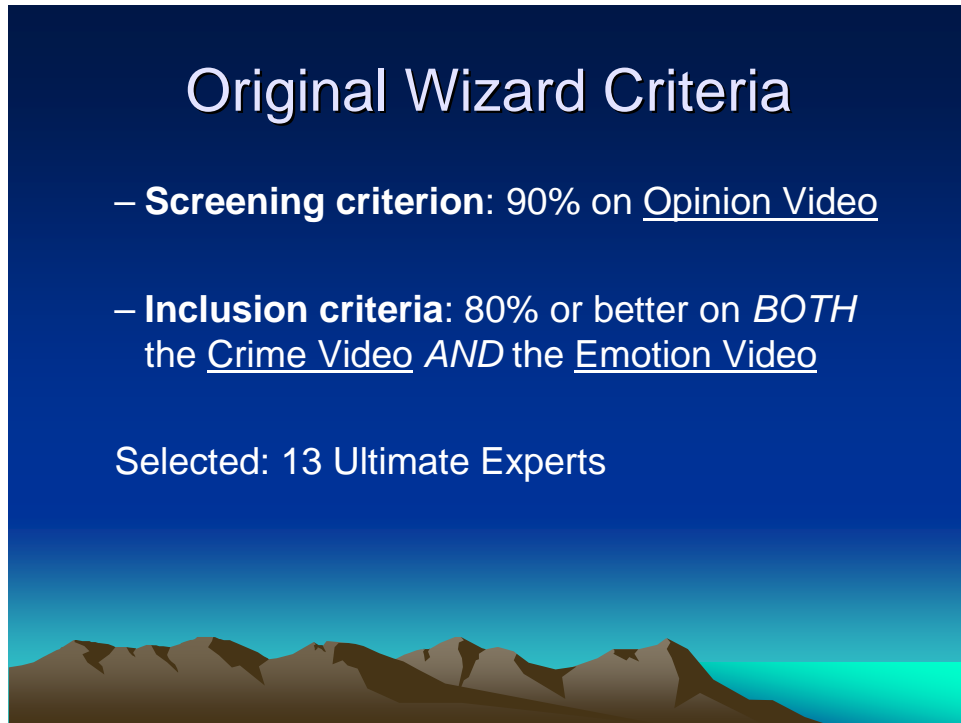
Definitions

- *Social intelligence:* "….the ability to understand and manage men and women, boys and girls-to act wisely in human relations."
- Thorndike, 1920, p. 228

- *Behavioral cognition:* "…the ability to understand the thoughts, feelings and intentions of other people insofar as they are manifested in discernible behavior."
- O'Sullivan & Guilford, 1975 p. 256

- *Emotional intelligence:*"…the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions"
- Salovey & Mayer, 1990, p. 189

One of the other contributions of Sternberg's theory is its emphasis of the function of intelligence, namely that being intelligent allows one to adapt to be successful in achieving one's goals. This emphasis on measuring functional abilities rather than theoretical factors has marked the last 30 years of emotional intelligence as Drew noted yesterday.

What I want to talk with you about today is a program of research on individual differences in the ability to detect deception. Detecting deception accurately involves understanding the emotional and personality baseline of the other person, and being sensitive to emotional and cognitive deviations from that baseline. It also involves the willingness to label a person as deceptive. This is a motivational and personal attribute that has been little studied. As such I believe that the ability to detect deception is one kind of social emotional intelligence.

This research program was started more than 30 years ago by Paul Ekman and Wally Friesen, who set out to study high states lies, because they were ones that could cause emotions in liars or truth tellers. They chose scenarios and arranged rewards and punishment to exacerbate the emotionality of the situation. The first measure of lie detection they devised was a video tape test of ten female nursing students, lying or telling the truth about whether they are watching nature films, or gruesome surgical films. The young women were motivated by their belief that the ability to control their emotion was related to their ability to be good nurses, which in fact it was.

Definitions

- *Social intelligence:* "….the ability to understand and manage men and women, boys and girls-to act wisely in human relations."
- Thorndike, 1920, p. 228

- *Behavioral cognition:* "…the ability to understand the thoughts, feelings and intentions of other people insofar as they are manifested in discernible behavior."
- O'Sullivan & Guilford, 1975 p. 256

- *Emotional intelligence:*"…the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions"
- Salovey & Mayer, 1990, p. 189

Let's call this the emotion video. Unlike many emotional intelligence tests in which the criterion for the correct answer is arguable (it could be expert opinion or subject consensus or muscle coding), the deception detection task has an unassailable accuracy criterion. What film the women are watching is a matter of fact.

Ekman, Friesen and I administered the emotion video to many law enforcement and other professional groups as well as hundreds of college students. We were quite surprised to find that every group we tested obtained average scores that were no better than chance; that is with 2 choices, is this person lying or is this person telling the truth, the odds that you are going to get one right is 50%, and 50% is what we found was the average score in every group that we tested. This finding of chance accuracy was found not only in our studies but those of other deception researchers as well. They reported the same thing. And how often do you find everybody finding the same thing? So another one of our knowns is that most people cannot tell when other people are lying, beyond what would be expected by chance.

Another finding from this research is that most people did not know how accurate they were in detecting deception. The correlation between their assessment of their skill at detecting deception, before they took the test, and their score on the test was zero. Other deception researchers have also reported the same thing, as has Patterson recently with respect to peoples assessment of their ability to judge their personality of others. So this is another known: most peoples' opinions about their ability to judge others are unrelated to reality. This raises questions about the meaning of self report measures of emotional intelligence. If such scores correlate with successful functioning in life situations that is very useful, and self-reports can be used to make such predictions.

And the question is why is this so? Are these differences due to motivational differences, differences in interests in thinking about emotions, some kind of social ability? This is another one of our unknowns. We do not know what those scores mean. So back to the situation in which we are testing all of these people and we are not finding any group where the average score is better than 50% or chance.

Then in 1991 Ekman and I described a group of Secret Service agents who, as a group, did significantly better than chance on the emotion video. In 1999, Ekman, Mark Frank and I described three other groups who were significantly above chance in their ability to detect deception. So we now have identified four groups of highly accurate lie detectors, highly emotionally intelligent people with respect to this particular task. We have found that although most groups do not perform well as groups within almost all groups there are one or two people who are highly accurate, but since these individuals occurred only about one percent of the time and had scored well on only a single deception detection measure, we did not pursue examining them.

Then two things happened that made possible the study that I will describe. Mark Frank and Paul Ekman developed two additional deception detection videos and Ekman was invited to give a keynote address to 1200 therapists. So rather than many small groups, we had one humongous group to test.

Definitions

- *Social intelligence:* "….the ability to understand and manage men and women, boys and girls-to act wisely in human relations."
- Thorndike, 1920, p. 228

- *Behavioral cognition:* "…the ability to understand the thoughts, feelings and intentions of other people insofar as they are manifested in discernible behavior."
- O'Sullivan & Guilford, 1975 p. 256

- *Emotional intelligence:*"…the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions"
- Salovey & Mayer, 1990, p. 189

And now there were two additional lie detection measures. In one of them the crime video, young men lied or told the truth about whether they had stolen 50 dollars. If they lied or told the truth successfully, that is, if they were believed, they received a financial reward. If they were not believed, whether they were lying or telling the truth they would not receive the reward, and they were told they would receive an additional punishment.

In the second deception detection task, the opinion video, young men lied or told the truth about a strongly held personal belief, such as whether cold blooded murderers should be executed. As part of his keynote address to the 1200 therapists, Ekman administered the 10 item opinion deception detection test video. He then gave the answers to the test to the therapists and asked those who had gotten scores of 90% or better who were interested in participating in a study of expert lie detectors, to write their names on the answer sheets. About 50 of the 1200 therapists received scores of 90% or better. By chance alone, 12 people should have scored so well, so we were hopeful that between twelve and fifty, there were some people who really were emotionally intelligent. Over the last eight years we have continued to lecture to various professional groups. As part of those presentations, Ekman, Frank and I have administered the opinion video tape. We asked people who received scores of 90% or better to indicate whether they would be willing to participate in a project on expert lie detectors. We have tested many groups as well as individuals who contact us through media coverage of our work.

Original Wizard Criteria

– **Screening criterion**: 90% on Opinion Video

– **Inclusion criteria**: 80% or better on *BOTH* the Crime Video *AND* the Emotion Video

Selected: 13 Ultimate Experts

On the screen is the original criteria that we used to qualify people for inclusion in what we are currently calling "The Wizards Project". After receiving scores of 90% on the opinion video and agreeing to participate, subjects are then given two other deception videos, the crime video and the emotion video that I described earlier. If they achieve scores of 80% on both of them, they are classified as wizards. We were calling them geniuses earlier but they are not geniuses in the ordinary academic sense. We started identifying wizards with our large group of therapists, but only 2 of the 50 that got scores of 90 % on the first test, qualified using what I'm calling the ultimate criteria. Then we noticed that some of the therapists obtained scores of 80% on one of the deception tests, but not on the other one, and the task they failed was almost always the crime video.

Odds of classification by chance

- 90% on **one** test
- 80% or better on **two** additional tests

- by chance alone:
  - 25 in 1,000,000

By this time we also had identified several potential law enforcement wizards: cops, judges, lawyers and arbitrators. And we noticed that they showed the opposite pattern, that is, if they achieved a score of 80% on only one of the two tests, it was always on the crime video that they succeeded and never on the emotion video.

So what you have here is a graphic illustration of what I have just told you. The law enforcement people did less well on the emotion video than the therapists. Almost all of the therapists did very well on the emotion video; only one of our 9 therapists did not do well on emotion, Conversely, the law enforcement people always scored 80% or above on the crime video but none of them scored 80% or above on the emotion video.

So we did a Chi-square analysis of these data, which was highly significant. So it seemed to us that what we were seeing here is the result of relevant learning experience, an illustration of tacit knowledge. That the therapists were good at emotion and good at opinion, which is what their profession entails, and not good at crime, makes sense. Recall, however, that we do have 13 ultimate experts, so we have 14 people who scored above 80% on all three tests. So we are going to be looking to see what kind of differences we can see, if any, among them.

This study is still underway. We've only interviewed about half of the experts so my report to you today is merely suggestive. We will be using the spouses or other close family member of each expert as their own control and we haven't done any statistical analyses on these data. But I will report the result of our interviews with some of these experts, in which we interviewed them about their lie detection skills, using a think aloud procedure.

I was thinking as I was listening to the two psychometric talks that preceded me that I'm a psychometrician and I'm going to be drummed out of the corp because what I'm going to be reporting is essential a clinical study at this point; we call it an ideographic study to make it sound more scientific.

What we do after the experts have been qualified, is that we play the videos and ask them to think aloud about what they are thinking, seeing, feeling, what is going through their mind, what are they paying attention to. I tape record this interview. This process takes from 2 to 4 hours. After that we give them a semi-standardized interview about their life experiences, their education, hobbies, work and all that kind of thing.

I also found with some of the groups, for example, the cops that they would not be very forth coming about their exploits, They seemed very uncomfortable being self reflective and were inhibited in talking about themselves. With them I found that if I got a group of these guys together or get them together with their friends, have some drinks in a bar, and just kind of talk, they would be much more open and I would get many more stories that would be more revealing of their personal characteristics. I also have tried with some of our other experts to socialize with them, to have them and their spouse to dinner, to have people out to a restaurant, so I can get to have a more of a feeling of them as a person.

Final Wizard Criteria

- **ULTIMATE EXPERTS** (n=14)
- 90% on the screening opinion task
- 80% on BOTH crime and emotion task

- **PENULTIMATE EXPERTS** (n=15)
- 90% on the screening opinion task
- 80% on EITHER crime OR emotion task

- Total wizards: 29

We will, however, be giving them personality tests of various sorts. I have delayed doing that because I want to get the perfect subset and as we heard from Gerard's talk, it is very difficult to know exactly what the perfect measure is, but we will be doing that. Now these 29 wizards have been culled from tests of over 12,000 people, so they are an extremely rare group of people. We are conceptualizing this lie detection ability as an essentially normal distribution. And we are looking at people at the far upper end of the distribution. So what have we observed about the people that we have been studying?

Wizard Lie Detectors Similarities

- Sensitivity to nonverbal cues
- Sensitivity to discrepancies
- Tolerates ambiguity
- Delays decisions
- Highly motivated
- Habitual practice
- Baseline accuracy, both behavior and personality

The first thing that we observed in the talk aloud procedure is that all of them are extremely sensitive to non verbal behaviors. They all talk about that. However, none of them uses the exact same non verbal behavior; some of them pay attention to facial expressions, but only the eyes; others pay attention to vocal intonations; other pay attention to how the liars and truthtellers move their heads. And they come up with and mention behaviors that we have not studied in 30 years of looking at non verbal behavior.

They are also paying attention to discrepancies. They will say "Well, his face looked like he's relaxed but his voice got very soft" (that soft voice that Paul talked about earlier). In order to be sensitive to and to process these kinds of discrepancies, these individuals have to have a very great tolerance for ambiguity. They will notice these things and it doesn't bother them that it doesn't fit together. The few spouses that I've talked with will almost consciously correct ambiguities, saying "Well he does such and such, but that can't be right, so I'm going to go back to the first thing."

Wizard Lie Detectors Similarities

- Sensitivity to nonverbal cues
- Sensitivity to discrepancies
- Tolerates ambiguity
- Delays decisions
- Highly motivated
- Habitual practice
- Baseline accuracy, both behavior and personality

In order to tolerate ambiguity and come up with the right answer, many of these individuals take a long time to reach their decision. I was quite interested in the fact that all of them had already taken the test before I sat down to interview them, None the less they wanted to go through and pay attention to each item before they would say what it was they were thinking. And most of them were markedly different (from their spouses) in their ability to remember what they had thought and to give very detailed explanations of them.

These individuals are extraordinarily highly motivated to do well at this task. They care about this task. Think about how we got identified "the wizards." We found them because they were attending a workshop on interrogation, improving your interrogation skills, or a workshop on learning how to detect deception. And we said "Okay you did very well. Are you willing to participate in a study/" Then they had to send me a consent form. They had to take two more tests and return the video tapes to me. They had to agree to be interviewed. So they were motivated to learn as much as they can about this. Another example of the degree of their motivation in this area: one of our early wizards calls me at 10 o'clock at night and he says: "There's an American Justice show on TV and it just started (it was like 1 or 2 minutes into the show) I think that man is telling the truth, and I want you to know that before the show is on that I think he is telling the truth and if I'm wrong I want to call you back and talk to you about why I was wrong." This is 10 o'clock at night, This is a man who has been retired from his law enforcement career for 5 years. So they are tremendously motivated to get it right.

Wizard Lie Detectors Similarities

- Sensitivity to nonverbal cues
- Sensitivity to discrepancies
- Tolerates ambiguity
- Delays decisions
- Highly motivated
- Habitual practice
- Baseline accuracy, both behavior and personality

Paul talked earlier about how difficult it would be, to widen the gap between the spark and the flame, the impulse and the action, and that it takes habitual practice even for a man like the Dalai Lama. Paul reported that he can't go 3 days without losing that ability. These individuals, and there is a range among them, only a few are fanatically driven, and are doing this habitual practice all the time. But all of them care. It's a part of who they are and they are interested in it, and they practice it.

The last thing is a bigger category. I have been quite impressed with these individuals with how they construe and understand the personality of other people. Yesterday, Karl Heider said that the Minangkabau can respond in terms of our categories for emotion, but that's not the way they ordinarily do, and this was implied in Gerard's talk as well, that if you use indigenous language structure you may get a different factor. And what I'm seeing with these expert lie detectors is that they understand discrepant non verbal and verbal behavior in the light of their understanding of who that person is but they use personality templates that are very different from our extroversion, introversion, conscientious, etc. They don't talk about those kinds of traits. It's not that they couldn't, but they have what seems to me fairly idiosyncratic ways of understanding other people.

Wizard Lie Detectors Similarities

- Sensitivity to nonverbal cues
- Sensitivity to discrepancies
- Tolerates ambiguity
- Delays decisions
- Highly motivated
- Habitual practice
- Baseline accuracy, both behavior and personality

One man, for example, said "He looks like a choir boy who hasn't been molested". What that implies, you see, is that he has a template of what a choir boy who has been molested is. And what that also implies is that he has seen such people. Another man described an African American man with Rastafarian hair curls and a nose ring, as looking and acting like a second grade girl. And emotionally, that's who the guy was. If you paid attention to the Rastafarian hair and the nose ring, you would be totally off base. But this astute man could see beyond the hair and the nose ring to get a sense of the emotionality of this person. Using that as his baseline, he was able to accurately assess deviation from that baseline in behavior and non verbal discrepancies and so forth. I am calling this the Miss Marple effect, from the Agatha Christie character who solves crimes by comparing the personalities of the people involved in a crime she is dealing with to the personalities of the people from her home town. And obviously the way in which she construes those personality characteristics are very different than we psychologists would.

Wizard Lie Detectors Differences

- Gender
- Childhood Trauma
- Verbal sensitivity
- Education
- Intelligence
- Profession
- Sensory acuity
- Role-playing ability

So far I have reviewed ways in which the wizards are similar to one another. This slide shows some of the ways in which they may differ among themselves and from other people. After we do the statistical analyses, we may find that more of the wizards than the control group have verbal sensitivity, but at this point what I'm noticing is that among the wizards they all pay attention to non verbal sensitivity but they differ in terms of how much emphasis they give to the nuances of language. A number of our experts are lawyers and they were particularly astute in terms of changes in the level of words that people would use, whether they would speaking at a very highly educated, articulate level and all of a sudden use a very inappropriate word, or would try to use a very fancy word when their ordinary level of speech was not at that level. But not all of the wizards do that.

The first variable on the list is gender. The non verbal sensitivity literature says that women are more sensitive than men in the ordinary kind of lab work that we do. In the expert lie detection work, however we are finding that it is really unpredictable at this point. We find more male therapists than you would expect by chance, given the sample we had; we are finding more female cops than you would expect by chance, given the samples we had. Gender seems not to be predictive of the extremely high level of accuracy in lie detection that we are studying.

Wizard Lie Detectors
Differences

- Gender
- Childhood Trauma
- Verbal sensitivity
- Education
- Intelligence
- Profession
- Sensory acuity
- Role-playing ability

The issue of childhood trauma is also still up in the air. Among our early wizards, many of them reported abusive childhood situations, parents who were either full of rage or alcohol, alcoholics or they beat them, so they reported a life time of having this kind of emotional scanning going on, where they were constantly aware of shifts in the emotionality of the people around them. But more recently our other experts are not reporting that. Now I am still continuing to probe for instances of early childhood trauma, but so far I'm getting mother and apple pie and everything was just fine and they were just people who were very curious and this was just one of many areas about which they were very curious.

In terms of education everybody in our sample has at least a college degree, because many of the cops, even if they only had high school degrees when they entered the police corps, usually would get further educations later. In terms of intelligence, there is a range. I think there are some that are just a little over a hundred IQ and there are some who are PhD, MD, PhD/JD's, extremely, extremely brilliant. So in terms of verbal intelligence there is a range.

Wizard Lie Detectors
Differences

- Gender
- Childhood Trauma
- Verbal sensitivity
- Education
- Intelligence
- Profession
- Sensory acuity
- Role-playing ability

The professions are variable. We have the large group of therapists I described earlier. In the final sample, we have 5 therapists out of the 29 wizards. We have a number of law enforcement personnel, local cops and federal agents. We have many arbitrators or labor lawyers and dispute mediators. Those were the groups that gave us the largest yield. So people who are in a situation where they don't have a side, where they are trying to make a decision between the employer and the worker and if they don't get a reputation for being even handed and fair they are not going to continue to be able to work. So that group had a very high yield rate. Dispute mediators also have a very high yield rate. Other differences may include sensory acuity. Many of the experts of the wizards report stories or others would tell stories about their ability to scan an environment, not only for social or emotional information; their scanning was for all kinds of visual sensory details; one of them is a bird watcher, another one has expertise as a hunter and is well known for his ability to trap in nature. So this kind of sensory acuity is something we will be looking at.

The other thing that some, although not all of the wizards report, is role playing ability. Wizards who seem fairly quiet and unobtrusive will tell stories or stories will be told about them of their working undercover, of their assuming characteristics totally unlike their everyday life, so many of them have this capacity. But not all of them do.

Wizard Lie Detectors
Differences

- Gender
- Childhood Trauma
- Verbal sensitivity
- Education
- Intelligence
- Profession
- Sensory acuity
- Role-playing ability

Another difference, that is not on here, but is in terms of the personality of these expert lie detectors. Many of them are introverts, and you would suspect that perhaps using our procedure that there might be a bias for that, since how you get qualified is that you are willing to watch a video tape and make an assessment of somebody else doing an interview of somebody. However there are 2 or 3 of the wizards who are flaming extroverts. They are not the quiet shy, observant types that some of the others are, but alpha personalities. Also some of these people who are very quiet may not be introverts. One wizard said to me "I'm quiet, but I'm not shy.", so he lays back and when he's ready and he is able to do the work, then he does it. Many of the wizards are over 40, but we have one law student who is 25. Most of the other wizards, I think, have this tremendous understanding of people's personality because their lives have been extraordinarily broad, deep and rich. One man is an arbitrator, has his blue suit on, but over the course of his life he has worked in the Appalachians with wild cat coal mine operators, he's worked on tug boats on the great lakes, he's worked with the lowest level of workers and now he is working with CEO's of multinational corporations. So he's just had a life experience, for example, that I have not had. He's seen and worked with all kinds of people, and obviously many of the law enforcement people, also have this broad array of experience with people.

Wizard Lie Detectors Differences

- Gender
- Childhood Trauma
- Verbal sensitivity
- Education
- Intelligence
- Profession
- Sensory acuity
- Role-playing ability

And that's why I think emotional intelligence does not exist in a vacuum. Moshe talked about display rules, and Paul talked about subtle expressions and leakage, micro momentary, so it's not enough to really recognize full facial expressions. That's easy. But when people moderate those expressions, for cultural or personal reasons like trying to appease somebody, that's the subtlety that, I believe, reveals emotional intelligence. And you need many kinds of skills, motivations, and personal attributes in order to do that. To my mind, it's like training Olympic athletes, you need to have basic talent, you need to get feedback, you have to have motivation on the part of the person that wants to become the Olympic athlete, and that potential athlete has to practice, practice, practice and that is essentially these what wizards of lie deception do. Thank you very much.

DISCUSSION

DR. KYLLONEN: Questions?

DR. GADE: What about actors?

DR. O'SULLIVAN: Actors I think will vary, I think they are in a business presenting a performance and we have not found any actors who are particularly good. We have one artist, who did painting of facial expression of emotion, but I think you have to be interested in what other people are feeling, not in what you, yourself are feeling, so I wouldn't expect a high yield rate from actors.

DR. STRICKER: Maureen, after you mentioned that this secret service group did above chance, you said there were three other groups. What were these three other groups?

DISCUSSION

DR. O'SULLIVAN: The other groups were selected federal investigators, if we just test the CIA forget it, they're 50%, FBI 50%, but there was a workshop where people in each of these agencies, FBI, CIA, and all that, were asked to send their best interrogators to an advanced interrogation course and that's who they were, they were people who in their agencies, people knew were good, at understanding others, the other group was a group of forensic psychologists who took three days to study this testing deception, so they gave up private practice days, and they gave up their weekends to come to this three day course so there's that high motivation which we referred to and the other is a group of judges, again who was interested to take the course, so that's who they were. None of those people are in this group, because we didn't collect this kind of data then.

DR. LEN WHITE: I recall reading a study a few years ago, where customs officials blew away all the other professionals. I'm thinking that in terms of related to an implication that people and professionals where they have opportunities to observe people who might be deceiving and also get feedback on whether they were right or not, might be an important thing. Now have you considered asking people things like how many people to you see a day, where you are trying to decide whether they're deceiving or not and do you get feed back on that?

DISCUSSION

DR. O'SULLIVAN: I think you are absolutely right. In fact Paul's article was about and I have an article out 2003 and PSPB and I have something like 17 reasons why we can't detect deception. But I think that one of the major ones is the fact that we don't get good feedback like we'll say I can always tell when my son is lying. But of course we only know the lies we caught him in. You know, we don't know what a good lie looks like, we don't what a good lie looks like, we only know what a bad like looks like. So I think the lack of feedback is a very important thing, and the other thing about the custom agents is that they don't have a side in a way, you know like a copy doesn't arrest somebody to be mean, they arrest them because they believe they believe they are guilty, and so that belief will color a whole bunch of things that they will see. And I think that is why the arbitrators and the dispute mediators much like the customs officials they may have a more balanced and open view about it.

DR. SIGEL: Do you think that there is any thing in terms of the flexibility these people have to make these judgments and whether they are open to begin with?

DISCUSSION

　　DR. O'SULLIVAN: Yes absolutely, and that would be a part of the fact that these broad and rich lives that they've had. They have sought that kind of diversity in their life, and they will look at somebody, like there is one guy everybody gets wrong on the opinion video cause he looks like a total goofball. And people will look, they will make this instantaneous attributional judgment and as we know, once you've made those judgments, it's almost impossible to change it, but these people don't make that judgment, they delay, and they say, well he looks a little goofy, what's he's doing in this study, okay and they keep it in mind. And what's interesting is these people like in terms of isms they are in their personal lives, many of them are extraordinarily bigoted and they will say outrageous things, that no academic would ever hear said. But when it comes to judging an individual person that's out the door, they pay no attention to what they're, so they are able to decouple and to be extraordinarily flexible in that way.

　　DR. SIGEL: But, they are also able to delay being reflective. Is that what you're saying?

DISCUSSION

DR. O'SULLIVAN: Yes, and I have one woman, this artist who I could see learning as she was doing the test. Because she noticed a particularly kind of behavior with the first man who, on the basis of other evidence she thought was telling the truth and then she saw the same thing, later on in about a man about who she was more equivocal and so she was essentially teaching herself as she was going through the tape. And that is tacit knowledge ability that I think is so interesting; where as, I've only done a couple of the spouses and they are much more wanting to get closure, wanting to make the decision of a lie or telling the truth. Where as the attitude of these experts are what is this wonderful thing I have before me? From what garden does this flower grow?

PARTICIPANT: I was interested in your comment about being bigoted, could you elaborate a little bit, because I have a follow up question.

DR. O'SULLIVAN: We have the N word, how women are so stupid, and I'm right there and they tell me that.
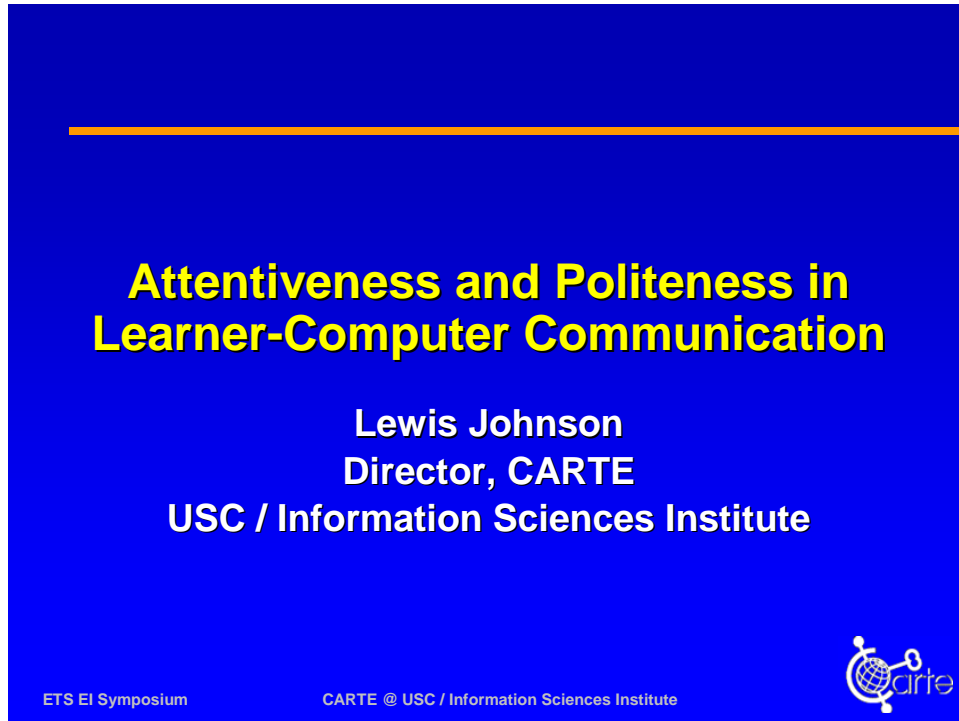
PARTICIPANT: They want to categorize, I was interested in taxonomy because you said the bird watching and some other things you said, intrigued me about the need to have a taxonomy of these various categories of nature, or other stimuli that they use the taxonomy to help understand their environment.

DISCUSSION

DR. O'SULLIVAN: I think the bird watcher was more about his interest in subtleties of sensory experience and he bird watched more for the pleasure of differences among them, and that kind of attitude of the pleasure of differences among them is more like what I would see. I don't see them looking to pigeon hole people, I think that's what we psychologist's do and I think that's why academic psychologists as a group gets 50% on these tests, you see, but when you are looking to find out how the people are different then you have a different mindset. Does that answer your question, we can talk later.

DR. KYLLONEN: I have a question while we're waiting for the next presentation to set up. Could you say a little bit more about the validation of these ideographic observations or findings?

DR. O'SULLIVAN: At this point it's very low; it's just how sensitive and emotionally intelligent you think I am. But what we are planning to do, however, is to use their spouses as controls, so we will content analyze the typed script of these videotaped interviews and do content analysis comparisons between the experts and the control. I also will be giving both the controls and these experts a variety of tasks on intelligence and what not. At this point, it's just to share with people an element of what an emotionally intelligent person might look like from a wide variety of professions. And what is so interesting is that these are lie detection measures and yet we're getting people who are from a whole variety of different fields and that's why we think there is some cross generality in the in the observations.

Attentiveness and Politeness in Learner-Computer Communication

Lewis Johnson
Director, CARTE
USC / Information Sciences Institute

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

ATTENTIVENESS AND POLITENESS IN LEARNER-COMPUTER COMMUNICATION

    DR. KYLLONEN: I would like to introduce our final speaker this morning, Professor Lewis Johnson from University of Southern California, Information Sciences Institute.

    DR. JOHNSON: I'm from a computer science background but I have been raiding the emotional intelligence literature quite a bit, as well as some of the other related disciplines. What I'm going to be talking about here is some of our work in modeling, attentiveness, and politeness in learner-computer communication.

The context of this is to create animated virtual tutors (we sometimes call them guidebots), that appear on a screen or in a virtual environment that are able to engage with learners in natural interactions to promote learning. We use these as guides, tutors, teammates, or as other story characters. This is relevant for education and training but this kind of technology is also applicable to commerce and entertainment applications. As an aside, I just want to say that in order to make these agents or guidebots work effectively, they have to have good emotional expression and often have to incorporate rich models of emotional processes. A lot of our work where this comes up is in what we call Interactive Pedagogical Dramas, which are dramatized scenarios where learners can understand how skills can be applied.

This shot here is from a project that we did called Carmen's Bright IDEAS. Stacy Marcella, Kate LeBoran and I fielded this at a number of cancer clinics around the United States. What it is intended to do is to help mothers of pediatric cancer patients learn to cope with their problems, many of which have a lot of emotional impact. There are a couple of aspects to this, one of which is to be able to create fictional characters that present the problem solving activity. In this case Carmen, the woman on the right, is a mother of a child with cancer. We did a lot of work looking at the emotion and the clinical psychology literature to create realistic expressions of emotion.

And then we created a virtual counselor, Gina, who responds to what Carmen says and does. Carmen, at any one time, can respond to Gina or to the learner who can influence what Carmen's emotional state is in the situation and to how she is coping with the situation. And to make this work effectively, this requires rich models of emotion including appraisal and coping which is realized in the Carmen character in particular. That is not what I'm going to be talking about today; I just sort of put that out as a teaser.

What I'm going to be talking about today is what technology is required to create more of these virtual Gina's, but which are able to interact with human learners in an effective way and which can operate in a range of learning settings, including more ordinary learning domains, such as math and science. Although we certainly recognize that emotional intelligence is important, the thing that we found to be of foremost importance is developing a model of social intelligence that we can incorporate into these agents. Primarily, the reason for that is to influence the motivation of the learners who are interacting with these guidebots. As I will explain, what we found is that motivational factors are things which good human tutors are very attentive to and we want to create virtual tutors that have similar capabilities. We refer to this as social intelligence because we regard the influences on the learner as taking place in a social context. Like Rosalind Picard, we are also influenced by the research of Reeds and Naps and others who liken human-computer communication to human-human communication. So, we actually found it useful to emulate aspects of social interaction in the agents as they interact with learners. What I am going to be talking about mainly today are two aspects of this that we have looked at in a fair bit of detail and think are important parts of this. One, which we call attentiveness, is the ability to observe closely what a learner is doing and detect potential problems before they become a serious motivational or emotional issue. The second is the role of politeness in interacting with the learner; basically how politeness can be exploited to have a motivational affect.

The context of this particular work that I'm going to be talking about is a project that we have doing for a while called the Virtual Factory Teaching System. This is an on-line simulation of factory processes and a number of universities have used this for teaching industrial engineering. It also is a topic for beginners who might potentially benefit from this, such as business majors, who find this difficult to master initially. Our goal is to be able to incorporate into the Virtual Factory Teaching System a guidebot that can monitor a learner's activities and provide help and feedback in a socially appropriate manner.

Experimental Study

- **Videotaped sessions of computer-based learning with human tutors, over multiple sessions**
- **Students read tutorial on line and perform series of exercises with Virtual Factory Teaching System**

Johnson, Pain, Shaw, et al: IUI '03, AIEd '03

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

Some initial experimental studies that we did form the basis of this work. We videotaped human tutors interacting with human learners in front of the Virtual Factory Teaching System. We scheduled multiple sessions so we could compare the interaction between the tutors and learners at later sessions as opposed to the initial sessions. The learners carried out a series of exercises and the tutors were instructed to provide them with assistance as they deemed it important.

Conclusions from Study

- **Wide variation in learners' preferred interaction styles**
    - Some prefer collaboration, some prefer working alone
- **Wide variation in confidence**
    - Between subjects
    - Over time
- **Tutor generally able to assess learner confidence, ability, preferred interaction style**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

One of the conclusions that we drew from this was that there was a wide variation in preferred interaction styles. One difference that we noted among the learners that we studied is that the way that they interacted with the tutor seemed to relate to whether they preferred to work on problems in a collaborative manner or preferred working alone; this would clearly seem to relate to personality measures that were talked about earlier this morning. We also noticed that there was a wide variation in confidence between the learners as well as a difference in confidence over time and that the human tutors were able to assess the learners' degree of confidence ability and preferred interaction style and then respond to that. So what we were trying to understand was how is the human tutor able to recognize these factors and then adapt accordingly?

Conclusions from Study (Cntd.)

- **Information used by tutor:**
  - Expectations from knowledge of task
  - Focus of attention (inferred from eye gaze, mouse location)
  - Confidence (inferred from learner questions, comments)
  - Effort expended
- **Focus on learner decisions and attitudes, rather than emotions**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

We looked closely at videotapes of the interaction and then followed this up with interviews of the tutors and these are some of the factors that we found to be most important. First, are the tutors' expectations derived from knowledge of the task. Basically, the tutor would look at what the student was doing or reading and from that, draw expectations of what the student ought to be doing next and if the student seemed to be at a point of indecision, the tutor might jump in. Focus of attention was extremely important which was inferred from eye gaze, where the student was looking on the screen or the mouse location. It was also clear that the tutor was trying to infer a degree of confidence; a lot of that was inferred from the kinds of questions and comments that the learners gave and the extent to which they seemed to be looking for confirmatory feedback from the tutor. Also, the amount of effort that was expended was important. As the student would go through these exercises the tutor would draw a general estimate of how heavily engaged the learner was from this pattern of interaction. So, these are some of the things we wanted to focus on and what we wanted to see was would it be possible for us to then create a guidebot that could draw similar inferences from similar sets of information. You see in some ways that these are similar objectives to some of the stuff that Rosalind Picard described, but we have chosen not to use intrusive measures such as galvanometers and things of that nature. We wanted to see how far we could get with the kind of cues that human tutors rely on, more than what a conventional desktop like what I am standing in front of here is able to respond to.
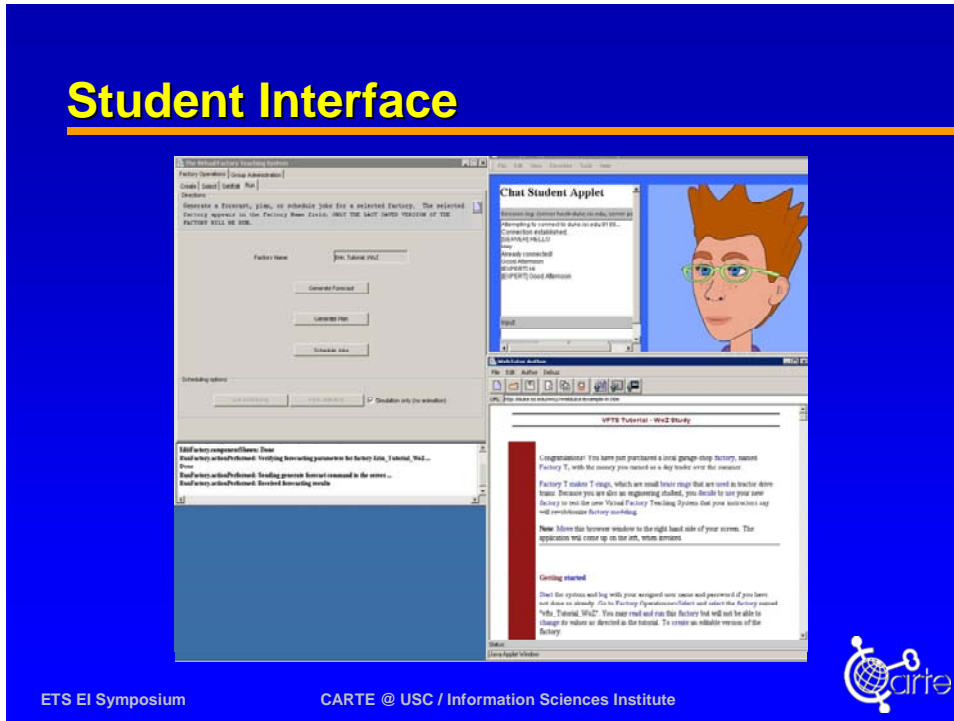
One of the key differences we wanted to achieve in terms of behavior was to create an agent that can proactively offer help. Again, we saw in Rosalind's talk the discussion of what some of the problems are when an agent jumps in with advice at an inappropriate time. How can we avoid that? That is clearly a dangerous thing to do if the agent has a poor understanding of what it is that the student is doing and why, so we need to be able to do a better job there. In order to support that, we created an interface that the student interacts with - here are the major components of it. On the bottom right is a tutorial window where the learner can read both information about the topic and instructions about what to do. The top left is the interface to the Virtual Factory simulation and the top right is the window for communication with the guidebot. The BFTS window and what we call the web tutor window are both instrumented, so as the student is interacting – scrolling, pressing buttons etc., all that information is being captured so it can be analyzed.
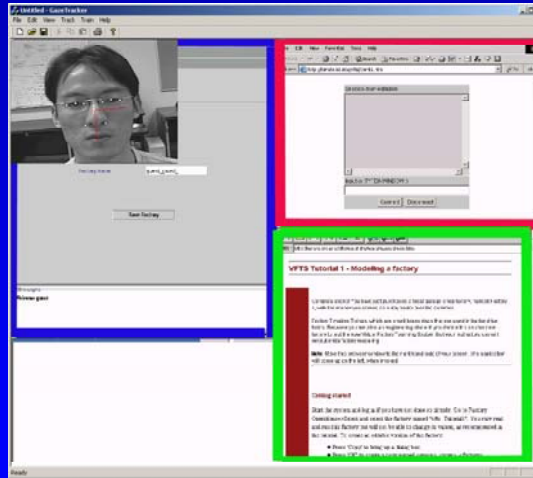
An important additional capability that we have been exploring is the use of gaze tracking to monitor what the learner is doing. Our goal here is to be able to do this in a robust way, without calibration and without special head stabilization hardware or anything similar to that; we want to make the effective use of gaze as one of the input modalities that the system can use. This particular system is based upon work by Larry Kite and his colleagues in the Laboratory for Computational Biological Vision at USC. Basically, it utilizes a low cost camera perched on the top of the display, which is focused on the learner's face and is looking for particular landmarks around the eyes, pupils, and the nose and then uses that to estimate head pose and direction of eye gaze.
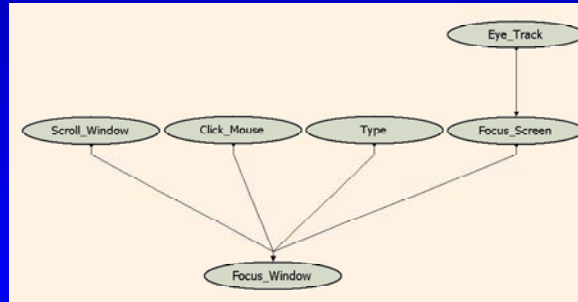
That works reasonably well, although without pre-calibration it's prone to error so we've been combining that with other sources of information that we gathered from the interface, as I mentioned, actions such as scrolling a window, clicking a mouse, typing, etc. These are all combined with the information from the eye tracking system in a dynamic Bayesian network. So when the student clicks on an object, we assume with high likelihood that the student is focusing on that object at that point in time. As time progresses, the likelihood that the student is still focusing on that object will decrease, but when we combine that with the information from the eye tracking system, that improves the overall accuracy of interaction. And what we've observed in our experiments is it actually does a pretty good job with the one proviso that there is a slight delay in recognizing when the student has focused from one new area of the screen to the other. But, it is adequate for what we are trying to do here, which is to model what we saw human tutors doing which was to be able to see what a learner's primary focus of activity is.
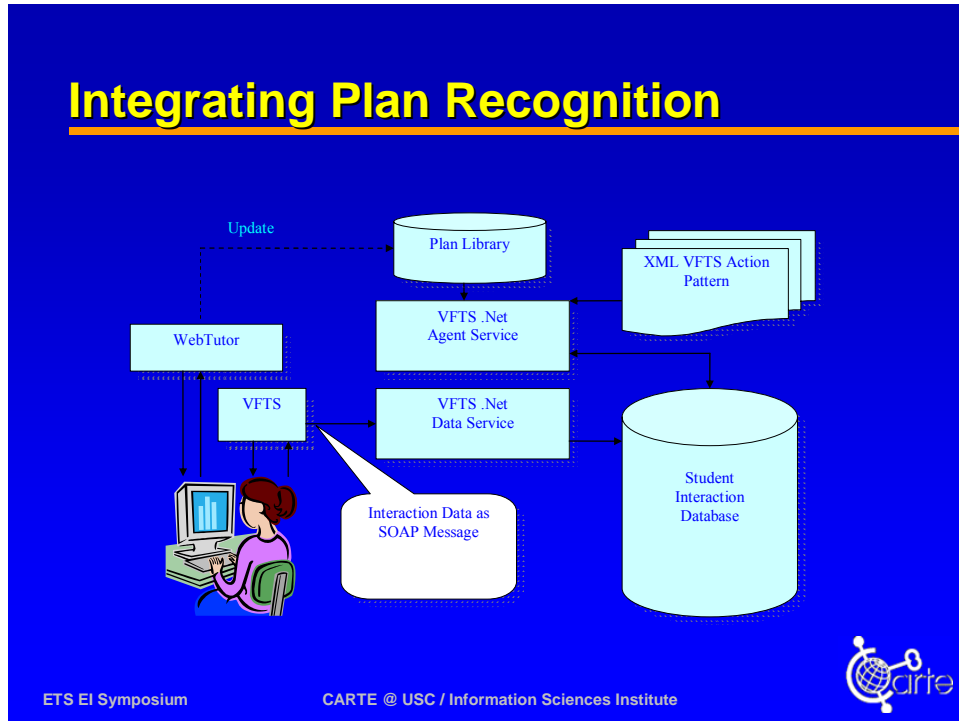
However, that by itself it is not enough. We need to combine that with some understanding of what the learner is doing in order to form expectations of what the learner is likely to be able to do next. What we combine that with is a plan recognition system which is basically taking this data from the user interaction and comparing that against patterns of different types of task solutions, which uses that to follow learner progress. Associated with those plans is not just the steps taken to perform but also the amount of time that we estimate is required to perform that task. We also incorporate into this estimation how much time it's going to take to read the instructions within the tutorial so that the system can draw expectations of when the student is likely to engage in actually carrying out the next step in the task. Given this time information, we can detect periods of indecision as well as incorrect actions by themselves.
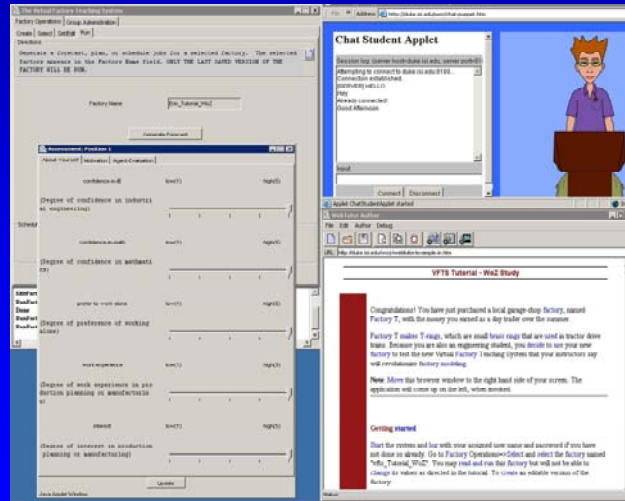
We also coupled this with a way that the learner can provide self-report of their own motivational state at various points in time. Colleagues at the University of Edinburgh have experimented with these types of interfaces and they found that they are reasonably successful. The factors that we are particularly concerned with here are: self confidence, sense of self control and sense of mastery of the subject matter area. Particularly, confidence and self control are the things that are the most likely to vary over the course of a session or a series of sessions and so these are ones that we want to try to track. So, we make use of the self-report interface but we want to make use of these other sources of information so that we don't have to rely upon self-report entirely.

Additional Inferred States

- **Effort**
- **Indecision**
- **Frustration**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

We are currently working on this and we see a number of things that we can derive from the data that we are collecting automatically. One is effort, which is an overall estimation of how much time the learner seems to be performing on reading the material and carrying out the task. Does it appear that the student is going though the material thoroughly or just glancing through it? We can detect instances of indecision at points where we expect that the student should be performing an action and we can tell that they are attending to the screen but they're not performing the action; there must be something that is causing them not to go forward. And then also, we see frustration as something that can potentially be derivable here by looking at instances of actions which fail to indicate progress and to suggestions where the student seems to be having difficulty accomplishing their goals. Finally, another important issue is to understand what the overall pattern or mode of work the learner is engaged in. That is, is the learner systematically going through and solving a problem, or is the learner just exploring the interface and the material? We don't want our guidebot to jump in with advice about what to do if the student is not really currently engaged on the task.

Again, this is something than can be derived by looking at the patterns of activity; that when we see that the learner is going through this process of reading material and carrying out actions and then reading some more, we can then provide feedback on that basis. But, if we see that they are carrying, for example, a series of actions on the simulation interface without reference to the tutorial materials, or vice-versa, those are indicators that a different mode of activity is taking place. So, those are the mechanisms that we are employing in order to be able to attend more closely to the learners' activities. It doesn't utilize the kind of sensors that Rosalind was talking about, but we think it is suitable for detecting a wide range of intervention points.

Interaction Tactics

■ **Once agent has determined when intervention is needed, it must determine how to intervene**

↗ **How to help the learner**
   • **Respecting learner's autonomy & sense of control**
↗ **How to influence the learner**
   • **Promote learner motivation**
↗ **Sensitive to social relationship with learner**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

Then the question is how to intervene when one of these intervention points has surfaced? For this, we have been trying to develop a model what we call Interaction Tactics, which is basically to determine how the system should intervene at a given point. We want to examine how to help the learner in a way that respects the learners' autonomy and sense of control, particularly if we know that there is some deficiency in the learners' sense of control. We also want to examine how to influence the learner to promote learner motivation in a manner which is sensitive to the social relationship between the guidebot and the learner. We are proceeding here with a hypothesis that such a relationship exists and it should serve as the basis for governing the interaction. This is something we are in the process of testing.

Examples

- **Tutor phrased comments in order to reinforce learner control and joint activity. E.g.:**
  - "Why don't you go ahead and read your tutorial factory"
  - "You want to save the factory"
  - "I'd skip this paragraph"
  - "So why don't we do that?"
- **Tutor avoided giving direct instructions**
  - Except for operating the interface

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

So, what do I mean by the different ways in which a tutor can influence the learner? Well, they contrast what I'm about to describe here. The typical feedback that you get from computer-based learning where you perform some action and it tells you, "No, that's wrong, try again" or "That's wrong here's the right answer", we saw in our videotapes very few instances of human tutors telling the student, "No you are wrong, do this." Instead, we saw a lot of phraseology like the following: "Why don't you go ahead and read your tutorial factory", or "You want to save the factory", or "I'd skip this paragraph if I were you", or "So why don't we do that." So these are all trying to convey suggestions or advice but they are not direct instructions of what to do. The only place where we found direct instructions were places like in operating the interface which said, "Click here in order to move forward." We wanted to try to develop a model that would enable our system to generate similar advice in a similar manner for similar reasons and, of course, understand what the real reasons are underlying this.

**Theoretical Framework:
Politeness (Brown & Levinson)**

- **Social actors motivated by face wants**
  - Negative face: freedom of action and freedom from imposition; autonomy
  - Positive face: consistent self-image, and desire that self-image is appreciated and approved of by others
- **Face-threatening acts pervasive in interaction**
  - Warnings, offers, promises, challenges, emotional displays
- **Face threat depends upon power, distance, ranking of threats due to social context**
- **Social actors employ politeness tactics to mitigate face threat**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

To develop this framework, we drew on the work of Brown and Levinson, particularly their model of politeness, where they looked at how polite communication takes place in a range of different cultures. Their model basically is as follows. Social actors, in general, are motivated by face wants, both negative face, that is the desire for autonomy and freedom of action and positive face, which is a positive self-image and a desire that that self-image is appreciated and approved of by others. In the course of interaction there are various kinds of interactions which can potentially threaten positive face or negative face or both. Warnings, offers, promises, challenges, emotional displays- these all potentially can be face-threatening acts and therefore people employ politeness tactics in order to mitigate or even avoid those face-threats. The extent to which those face-threat mitigation strategies are applied depends upon on the relative power relationship between the parties, the social distance, as well as the ranking of threats due to social contexts or culture.

Role of Politeness Factors in Tutorial Interaction

- **Common tutorial actions (criticism, advice, hints) are face-threatening acts**
- **Tactic failures impact agent's positive face**
- **Face threat depends upon distance**
  - ↗ **Distance depends on duration of interchange, established trust, learner's negative face wants (preference for autonomy vs. collaboration)**
- **Choose tactics to promote learner positive face, mitigate negative face threat**
  - ↗ **By promoting shared goals**
  - ↗ **By avoiding direct instructions**
  - ↗ **By reinforcing positive (self-)assessment of goal achievement**
  - ↗ **When dictated by social distance, learner motivational factors**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

What we found is that we could apply this notion to understanding the type of interactions that take place in a tutorial or coaching setting. First of all, recognize that criticism, advice and hints, are all potentially face- threatening acts. They threaten positive face by implying that the learner is not doing a good job at what they're doing and they are potentially threatening negative face by controlling what the learner does and giving them less autonomy. We also noted that it's possible that if a tactic fails, that would impact a speaker's positive face and this seemed to us to be something that is potentially relevant for agents that make mistakes and need to apologize to the user when that mistake occurs. We asked the question of whether the face-threat depends upon distance and we are still analyzing our human-human transcripts, but it does appear as if there is a tendency there, at least for interacting with some learners. That is to say, we noted that in second sessions of our studies, that there were many more instances of the tutor saying, "Okay no, do this now", much more of what Brown and Levinson would call bald-on-record interventions. This suggested that there is some change in the interaction over time which made the tutor be more comfortable in being more direct in their criticism. Also, we thought we could potentially build on a tactic that Brown and Levinson talk about, which is to promote positive face which then promotes the sharing of goals between the speaker and the hearer. This is what teachers have to do a lot of; they need to try to get the learners engaged in the activity that they are involved in and to share the goals of the teacher. So the question is, to what extent could the politeness model actually be a method for trying to promote that? I think to some extent it can.

Example Interaction Tactics

- **Rhetorical requests to give hints**
  - "Can I give you a hint? Try this…"
  - Question reinforces learner negative face; failure to wait for answer avoids positive face threat
- **Hints phrased as questions**
  - "Do you want to do *x*?"
  - Reinforces learner control (positive face), can influence learner goals (positive face)
- **Hints as suggestions**
  - "You could do *x*."
  - Similar face effects as questions

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

To give you a sense of how we apply the Brown and Levinson model in particular instances, I'll analyze a few of the tactics that we observed in our data. Here, for example, we observed a tutor saying, "Can I give you a hint? Try this". So, phrasing it as a question reinforces learner negative face because it's then up to the learner to either accept that hint or not. But notice that offering a hint is potentially face-threatening because then the other person has to decide whether or not to receive it. So, by going on and saying, "Well, try this", as we would analyze it, it's a way of reinforcing negative face from the advice but avoiding the negative face-threat of the offer. When hints are phrased as questions, we do similar analyses in terms of negative face. In terms of this group of learners, most of what was going on was having the tutor reinforce negative face on the part of the student, but we did see that these could potentially have a role in influencing the learner goals. For example, do you want to do" X" now as sort of suggesting that the learner should adopt that goal.

Interaction Tactics (Cntd.)

- **Hints as suggestions of joint goal**
  - ↗ **"Let's do *x*."**
  - ↗ **Suggestion mitigates negative face threat; reference to joint goal influences positive face wants; depends on learner autonomy preferences**
- **Hints as references to tutorial authors**
  - ↗ **Deflect blame for face threat to authors**
- **Imperative hints**
  - ↗ **Used only when blame is deflected (I.e., to interface), or possibly when distance is reduced**

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

In a more extreme case, like the suggestion of a joint goal, "Let's do this now", this is something which is addressing both negative face and positive face. This is one of the standard tactics that Brown and Levinson refer to, that one way of promoting positive face is to promote joint activity. The places where you would expect that there not to be these kinds of face-threatening acts are the places where there is no threat of a face-threat being made by the tutor. In fact, what we see is consistent with that, that what the tutors we studied did a lot of was say, "Okay, the tutorial says you need to do this now." Since the blame is on the tutorial, then the tutor can be more direct in terms of what it can say.
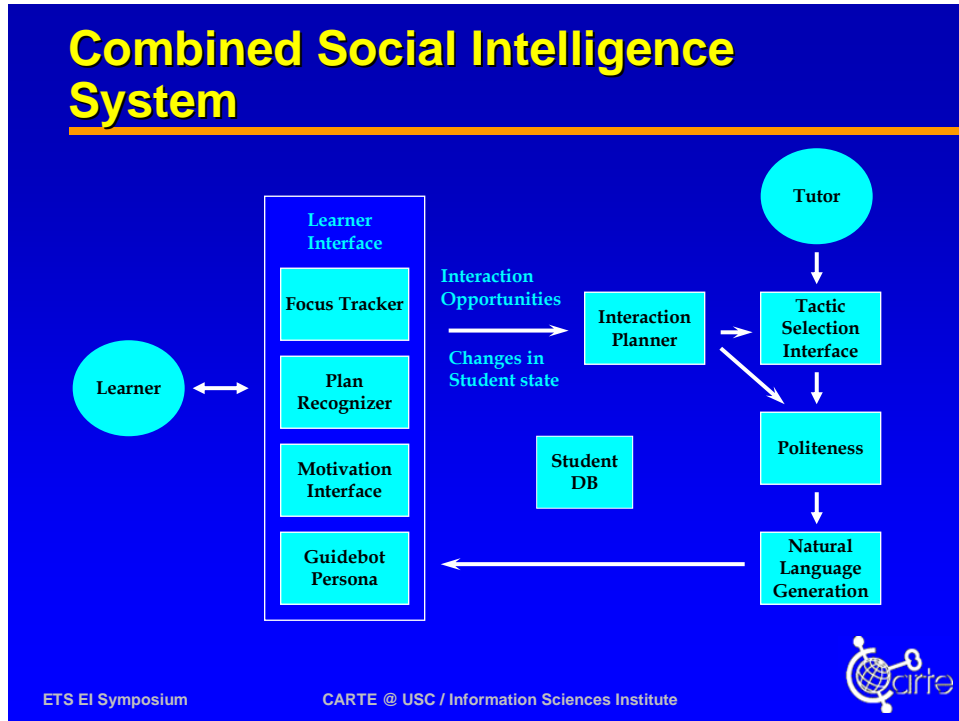
We now realize this in the form of a complete social intelligence system, which, on the left, has these abilities to track the learner's state, both at a cognitive and motivational level and also to detect interaction opportunities and report changes in a learner's state. Then, we have developed a tactic selection interface which selects an appropriate tactic. This can be used in two ways - we can use it under control of a human experimenter as well as under automated control or some combination of the two. This gives us some freedom in terms of what experimental approaches we want to take. The politeness module decides how to realize a particular tactic in English and is ultimately coupled with the appropriate non-verbal gestures. Finally, the natural language generation and the gesture commands go back to the guidebot persona to determine how it then interacts with the learner.

How do we implement the politeness model in this fashion? Actually, we had to extend the Brown and Levinson politeness model to some extent. We found that in order to be able to choose the appropriate strategy, we needed to take different categories of tactics and classify them both according to their potential negative face-threat and their potential positive face-threat and then choose the one which has the appropriate evaluation for each of these, according to the particular type of face-threatening action. For each type of face-threatening action, advice, etc., we would collect a set of tactics that could be performed and then the politeness model chooses the right one from among those. Now, this can be varied by the social distance that is specified. In other words, there is a parameter to this model that you can specify, which is a slider that you can control a greater or lesser social distance. As you reduce social distance you get more instances of bald-on-record advice. Finally, we provide a way of directly augmenting the desired influence on learner positive face and learner negative face. Again, this is an extension of the Brown and Levinson model; we're saying that we want this agent not to just be polite but by being polite, to reinforce learner autonomy or self-confidence. So, we realize that here by giving you the ability to augment the importance of negative face-threat or positive face-threat or both. Then, the tactic generator chooses from a library of utterance patterns that most closely matches the parameter setting.

This is work that we are currently doing and we've got a process of setting up a couple of evaluations here. One question that we are interested in is how much additional advantage do we get from this kind of attentive learner modeling activity versus a more reactive one? This gives us the potential for a larger number of intervention opportunities. We need to evaluate that both from the learners' perspective - were these appropriate interventions, and also compare this against what a human expert tutor would say. Again, we use this design in order to support that we can set this up so that the system is operating automatically but a tutor can be sitting in another room observing what recommendations the system is making in terms of intervention opportunities and commenting upon their appropriateness. As far as the politeness model goes, we certainly don't claim that this is a complete account of how social intelligence is manifested in computer-learner interaction, but we want to understand this particular mechanism in detail, mainly what effect that it could have. The way that we see to do this is to compare the version which I described, which is trying to deal with these face-threat issues versus a condition which ignores these, which freely criticizes the learner whenever they make a mistake and makes direct instructions of what to do. And then we would evaluate that, not just in terms of the learner performance and the learner attitude toward the system, but we also we want to understand what the impact on motivational state is. Are we right that this approach will actually result, after repeated interactions, in giving learners a greater sense of self-efficacy in their interaction? So I will stop there and entertain any questions.

DISCUSSION

DR. KYLLONEN: Questions?

MS. PLANTE: The reference you made to the difficulty and mastery of the system and its use currently in higher education programs. What are your plans to look at that from other perspectives and simplifying the systems such that they could be used in support of K12 learning?

DR. JOHNSON: I think that there are always things that could be done to improve the system but, I think our particular interest here is in understanding where individualized feedback and assistance would be advisable. I guess what I would answer is that we could certainly do this but the point of this is not to improve this particular system per say but to develop a set of techniques that could be applicable to a range of different learning applications in order to provide assistance to learners when they are getting started. We regard that as being a general problem that is worth trying to address.

DR. KYLLONEN: You looked at what the human tutors did and you captured all these activities or tactics. My question is: Did you make any attempt to try to sort out which of those activities were more effective on the human tutor's part? Or was the assumption that these were good tutors who employed effective tactics?

DR. JOHNSON: We did a lot of analysis of one tutor in particular who had received awards of excellence in teaching and so we presumed that these tactics would likely be a factor. But, we haven't done the comparison that you describe of course we can now easily go back and do that, as I said let's take out all of those tactics and then compare. Then, once we see the main effect, one of course can do further analyses and say what's more important here, to address negative face or positive face, or are there some other things we should be taking into account instead? So, that would be a topic of further investigation.

PARTICIPANT: How are you incorporating the gaze tracking and the other measures into the mentor's interactions?

DISCUSSION

ETS EI Symposium   CARTE @ USC / Information Sciences Institute

DR. JOHNSON: I neglected to mention before that I did include a working paper describing some of this work in more detail and I think that will be available to people who are interested in it. To answer your question, how is the gaze incorporated with the other sources of information? Each paragraph in the tutorial materials is marked as basically using xml, so the tutorial interface indicates what information is currently visible. The gaze tracking information is used to say, "Okay, the student is currently reading about topic X or now the student is currently looking at the screen." We couple that with the information of what actions the student is actually performing. What the gaze enables us to do is to detect situations where the student seems to be thinking about what to do next or seems to be unclear about what to do next, without having to wait for them to make a mistake; without requiring them for the system to react. We see this as being potential for aborting a lot of the frustration issues that we observe when students aren't sure what to do next or they do something wrong and they get the unexpected effect. We think that gaze will help a lot there.

PARTICIPANT: Does the mentor change the quality of his or her response on the basis of that or is it just the mentor either appears or doesn't appear?

DR. JOHNSON: Quality is primarily influenced by the internal model of learner self confidence and self control. Those two, as well as whether the student prefers to work collaboratively or not, all then can determine whether or not a particular intervention takes place and then also how that intervention takes place. If a model says that the student has a lower level of self-confidence, then basically the parameters of the politeness model, which promotes greater positive face, will be increased. So, you will see more interventions of the form, "Let's try that now", or "How about if we do this?" Things of that sort.

DR. LEN WHITE: Part way into the session, can a learner change their preference in terms how attentive how they want the guidebot to be?

DR. JOHNSON: Absolutely. And also with preferences in terms of, does the learner want to have this animated character at all or if the animation or speech be on or off. We use synthesized speech and that has a trade-off. On one hand, it is a way of providing feedback information to the learner without distracting them from their current activity, but on the other hand, you can only provide a limited amount of information before that becomes a tedious mode of delivery. We've noted that in some of our other studies that learners have different preferences in terms of what modalities that the guidebot interacting with them use.

DR. KYLLONEN: I was wondering if you could imagine a next generation version where you have video cameras looking for the micro-facial expressions. You recognize the look of contempt on the user's face and respond to that appropriately.

DR. JOHNSON: I will tell you some of the things that we're interested in getting from the face and they are not things that I've heard talked about today. One, are indicators of boredom and fatigue. The vision people that I've been talking with suggest that eye closure can be an indicator of that, so that's one thing we can potentially look that. And as far as the facial display issue goes, it's a different situation interacting with the computer from interacting with a person. What we observed from looking at our videotapes is that there are some learners who are much more expressive in their faces than others. However, I don't believe that by itself is going to be sufficient, there are definitely some learners that are working along and they're just not displaying very much. And it's not like they're lying, they're working, they're just not emoting, but I think that's fine. For the model that I described, that's one more source of information that could be combined with these other sources of information. I think what we're talking about is getting as many different sources of information as we can that are potentially uncertain and then try to combine those to get as accurate a moment-by-moment estimate of the learner's state as possible.

DR. KYLLONEN: That's it for this session. I'll hand this over now to Rich.

SESSION IV: ASSESSMENT
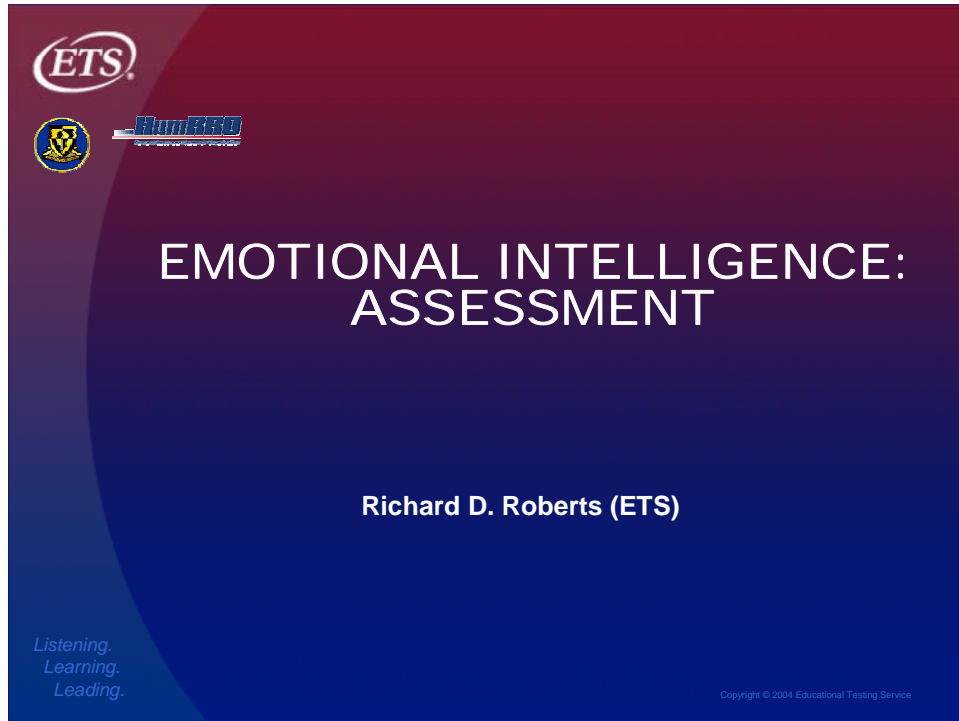
DR. ROBERTS: In this session, we have assembled a panel of participants who will discuss various measurement approaches to the assessment of emotional intelligence. Our first speaker is Faba Sala, from the Hay Group.

Assessing Emotional Intelligence Competencies

Fabio Sala
Hay Group,
McClelland Center for Research
November 14, 2003

HayGroup®

ASSESSING EMOTIONAL INTELLIGENCE COMPETENCIES

DR. SALA: Thank you very much; it's a pleasure to be here. Of course thanks to Rich, for all of his efforts in helping make it happen.

Let me begin by telling you a little bit about myself. Where I'm coming from: I work with the Hay Group, and we are a global human resource management consulting company. We help clients and businesses with a wide variety of business issues, problems and challenges. We essentially specialize and focus in on people and how to achieve results with people. I'm at the McClelland Center for research and innovation, named after our co-founder David McClelland. At the center there we are largely responsible for developing capability within the larger organization, so what that means is that I spend part of my time doing applied behavioral kinds of research, and also working directly with clients, applying what we've learned. Although we're here talking about emotional intelligence, it's a relatively small part of things that we do.

So let me first just explain where we are going to go today. Ultimately I'm going to discuss the Emotional Competence Inventory, which is the measure that we use to assess emotional intelligence competencies, but before that I just want to give you some historical context of how it was that instrument came to be developed. That will lead to a relatively detailed discussion of a behavioral event interviewing methodology, which Dave McClelland and others co-developed and we've been revising over the last 25 years. Of course we will get into some emotions within competencies, which came out of a lot of the work that we had been doing with this interview methodology. Then we'll get into the psychometrics of the ECI, which has been through a couple of iterations. We'll look at some correlations with work performance: we have a couple of correlational studies that look at the relationship between ratings on this assessment instrument and two really nice performance indicators in a couple of different settings. Then we'll sort of wrap of with the knowns and unknowns--I've highlighted a couple but there clearly are a lot.

One thing before moving on, in terms of assessing EI competencies we are really going to focus on the ECI, and we'll talk a little bit about the interviewing methodology. I think it's worth noting that we have a number of ways of trying to get at this thing and these are just two of them… other things like assessment centers putting people in actual situations that will possibly arouse emotion like challenging situations… and code them on a relatively rigorous checklist and then validate our assessment using that kind of criteria, as well as what's called shadowing--but that is really just walking around and watching real situations and following managers and executives and coding their behavior and trying to triangulate all these various types of assessments and figure out whether or not we really are capturing the essence of this person, asking whether what we're capturing makes any sense and has any utility.
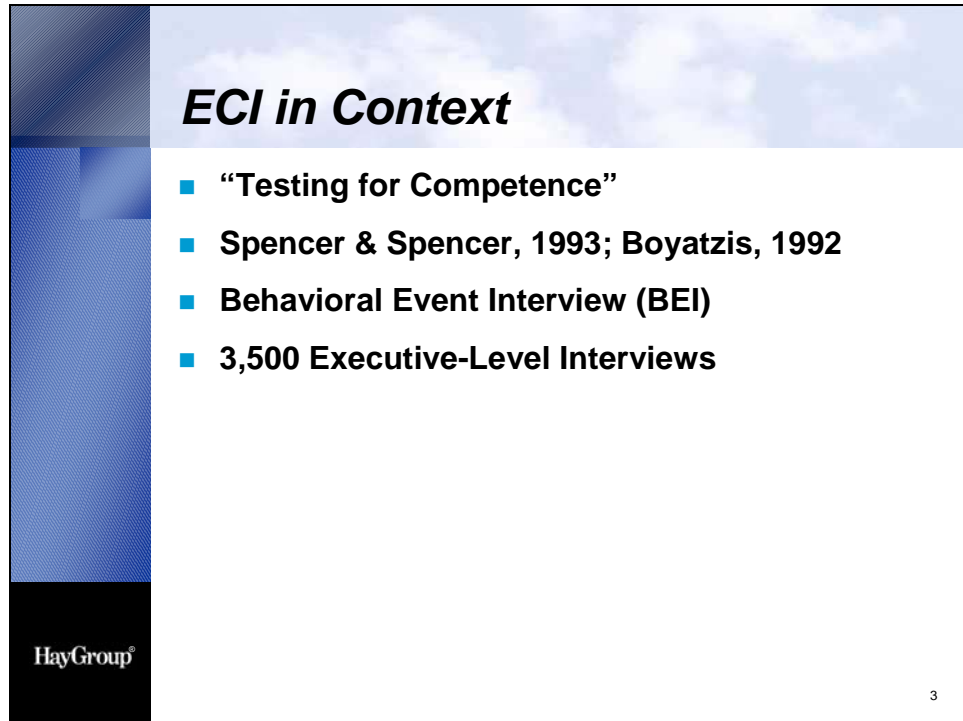
ECI in Context

- **"Testing for Competence"**
- **Spencer & Spencer, 1993; Boyatzis, 1992**
- **Behavioral Event Interview (BEI)**
- **3,500 Executive-Level Interviews**

HayGroup®

3

So let's get started. Ironically, the talk actually did start here; McClelland wrote an article in 1973 where he sort of urged folks to test for competence rather than intelligence. It's nice to be back here at ETS talking about the issue. Spencer and Spencer and Boyatzis are a couple of authors that have worked closely with the McClelland Center over a number of years and sort of taken some of the points that McClelland argued back in '73 and started to develop this Behavioral Event Interview methodology. Essentially, it's very similar to the plan-again critical incident assessment where you set down with someone and you ask them to tell stories about recent situations in the workplace--it could be anything. It could be about how they launched a new product, or about how they managed a subordinate, or anything. Then we take them through a pretty painstaking and rigorous process of trying to extract what actually happened. Ideally we like to actually go back there and be with them, and walk through this situation with them, observe them, stop them in the moment, as things are naturally unfolding and say, "what were you thinking there, why did you say what you just said?"
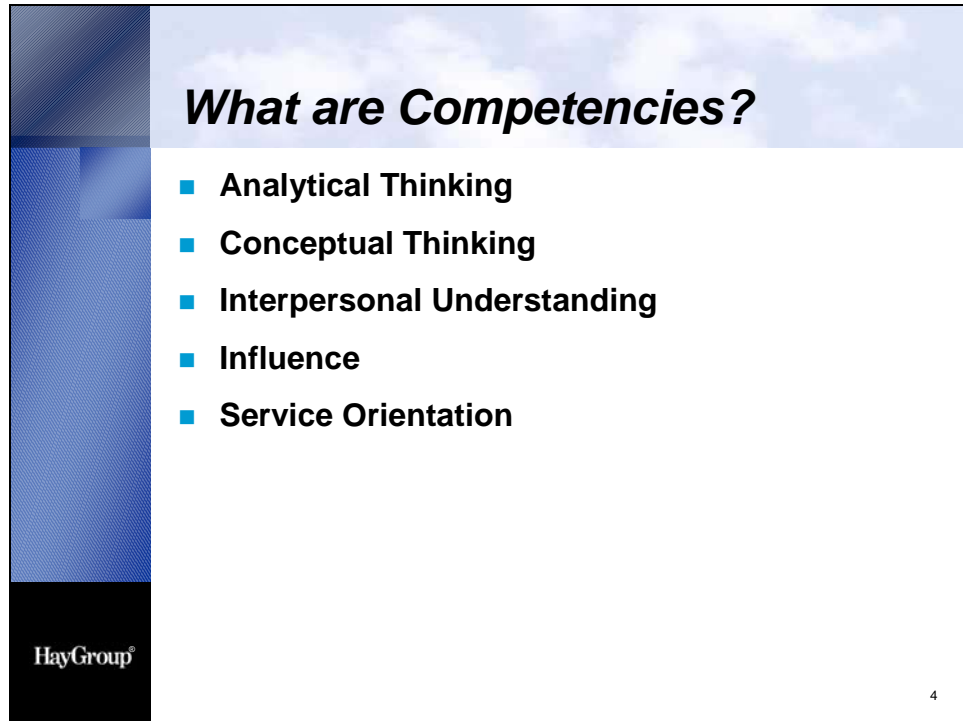
ECI in Context

- **"Testing for Competence"**
- **Spencer & Spencer, 1993; Boyatzis, 1992**
- **Behavioral Event Interview (BEI)**
- **3,500 Executive-Level Interviews**

HayGroup®

3

So that's what we try to do, that would be the ideal for us, but since we can't do that we ask people to recall and reconstruct the events and with a relatively detailed probing process, and we like to think we do a reasonably good job of capturing what actually did happen. And it takes some time to develop the interviewing skills to get to the level of sophistication that we believe makes the method credible. So we've done this with a lot of people--at this point we have over 3500 interviews with executives, mostly managers, but also lower level people in a wide variety of organizations, a wide variety of industries, pretty representative in terms of male and female. During the entire process I've been revising the method.

What are Competencies?

- **Analytical Thinking**
- **Conceptual Thinking**
- **Interpersonal Understanding**
- **Influence**
- **Service Orientation**

HayGroup®

4

This led to is a sample of what are called competencies. We have a sample of generic competencies and every time we do an analysis of, let's say, a group of people, it's always criterion based, so we have a sense of how effective that person is in their role--whether or not they're nominated as outstanding, or whether or not they're in some criterion group that indicates their effectiveness. So what we did after content analyzing all these interviews (which we can do reliably, by teaching people how to code), is glean from the narrative data what are called competencies--we will get to those in a minute--and identify a set of characteristics.

It's really a wide set of… it's personalities, it's skills, it's traits, it's cognitive abilities, and essentially to make the cut, to be a generic competency, they have to over a period of time be able to distinguish outstanding from average performance, and I mean that in a very global sense. If they show up, they have some values and utility for us; it's very much empirically driven, very much grounded in theory.

So there is some background. If you can imagine all this happening over a period of time, one of the things that seemed relatively clear is that these competencies seem, to a large extent, to be based in emotion, or at least part of what drives them is emotion-based, or what describes them, the words that are used to describe them, are words that you might use to describe emotion.

What are Competencies?

- **Analytical Thinking**
- **Conceptual Thinking**
- **Interpersonal Understanding**
- **Influence**
- **Service Orientation**

HayGroup®

5

Here are just a couple of examples of competencies that have been identified. Analytical and conceptual thinking are relatively straightforward and essentially, analytical thinking is kind of breaking things down, thinking about things systematically, thinking about how things are causally and linearly related, and then getting more sophisticated and seeing how there are multiple relationships between different things: A may cause B but only under condition D--that kind of reasoning and intellectual horse power.

Conceptual thinking: the sense that idea that people have this ability to take ideas or constructs or observations that are seemingly unrelated, and then be able to put them together in a way conceptually that makes sense. Maybe, for example, a real high level of this is to create a new concept or construct nobody has really every thought about or identified (maybe not nobody literally but something like that)--sort of piecing pieces of information together that make sense, and that are useful and helpful for people.

Interpersonal understanding is another competency that emerged from these interviews from these data, and I'll save that, and we'll use that one as sort of an illustration, what the competency looks like and feels like and essentially the extent to which you understand a person's inner life, another persons inner life--let's save that for a moment.

Influence is the ability to get people to be persuaded by an argument--say I'm here and I'm trying to persuade you that, what we have is credible, and I'll show you a couple of reasonable arguments, and then I'll leave it at that. Or maybe I'm in a meeting with you and I'm trying to get you to do [something], to adopt a perspective that I'd like you to take--maybe if I'm a little more sophisticated I might think about how it would also suit your needs, [how it would] benefit the both of us and tap into that, and maybe I could get other people to come in and persuade. So influence is another competency that we see, that emerges in this context and in these situations, and there are various levels of sophistication.

Service orientation is another competency, I was thinking about talking about that in detail, but I scrapped it. But essentially that's not really a personality trait. It really is more an orientation toward giving high quality service, and you can see this in a wide variety in contexts.

The competencies… really they sound like personality traits, and some of them are. Obviously there are also cognitive kinds of traits, certainly some skills… the extent to which they measure emotional intelligence as we'll try to define it, I will leave that up to you. Let's look at it in a little bit more detail.
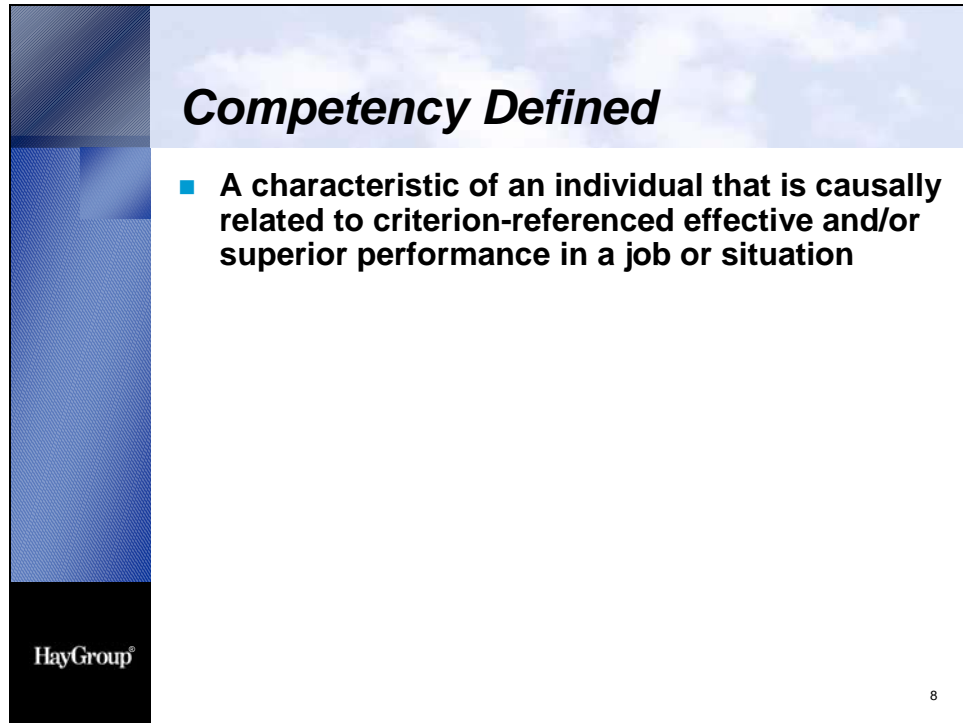
Competency – I have sort of I've been describing them a little bit, but here is our formal definition: "Characteristic of an individual that causally related to a criterion-referenced effective and/or superior performance in a job or situation." So really, all it is saying is that it is anything that potentially leads to outstanding performance in a job, or situation. When I say causally, I don't meant that in the literal sense--that we've done controlled experiments and carefully tried to demonstrate causality--that's not really what I mean there. But essentially it's through this process of identifying of what it is that distinguishes outstanding from average performance, and it could be essentially anything.

One of the interesting things that come out of this kind of analysis that we identify what we call uniques. And really, all that it is that we keep a very open mind when we are studying a sample. When we say, [for] managers in a particular organization or a particular industry we are not exactly sure what it is that leads to outstanding performance, we know, based on our experience, that there are certain qualities and characteristics that people tend to [have] which tend to make them more effective, but in a given environment we may not know exactly what they are, so we are open minded to these uniques. So, let's say we do 20 or 30 interviews and then we have various people blindly code them to get transcribed and code them and so forth, and people look for uniques or whatever--it might be an orientation toward planning or organizing a skill; it could be anything.
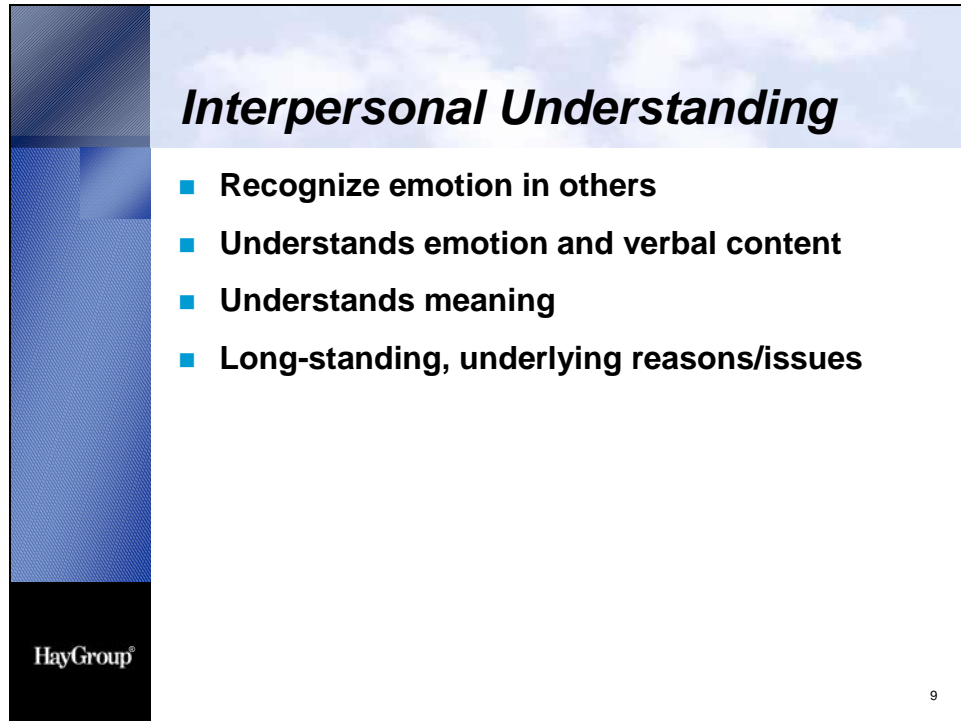
Competency Defined

- A characteristic of an individual that is causally related to criterion-referenced effective and/or superior performance in a job or situation

HayGroup®

8

And so if a pattern emerges and several different people say, hey, when I was reading an interview about this person, [they] seemed to be doing this a lot, and it seemed to be helpful--and if that is a consistent pattern that we see in that group, then it might be a competency that seems to distinguish outstanding from average performance. And if we see it frequently enough over time, over a number of contexts, then it makes the cut and it becomes a generic competency that we think adds value in a wide variety of contexts. I hope that gives you a sense of the concept of competencies, where they're from, what they mean, how they're used, how they were created.
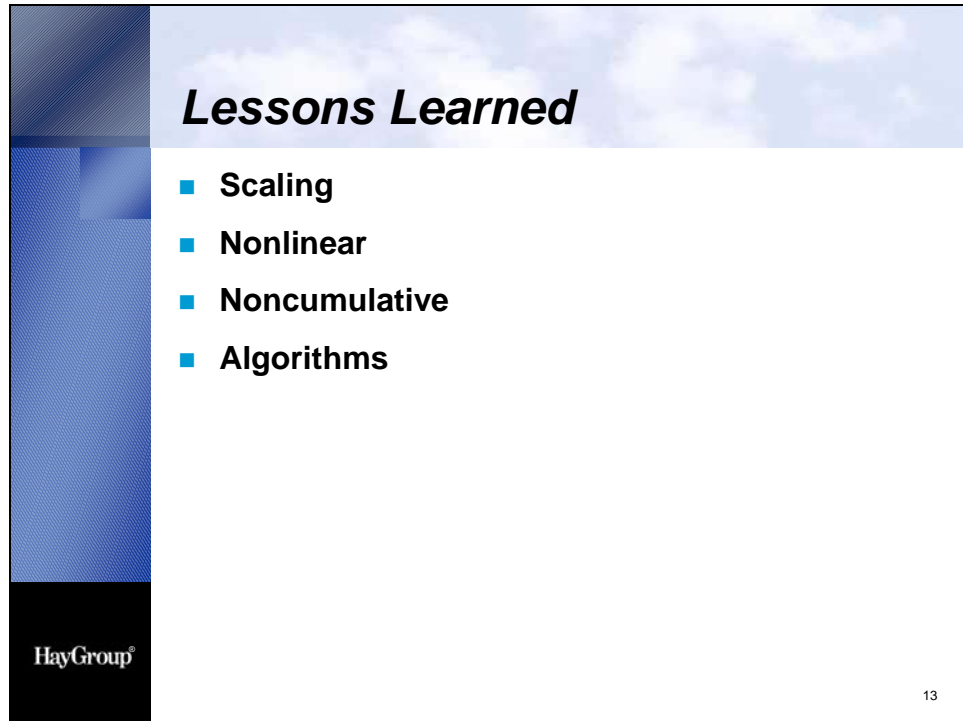
Interpersonal Understanding

- **Recognize emotion in others**
- **Understands emotion and verbal content**
- **Understands meaning**
- **Long-standing, underlying reasons/issues**
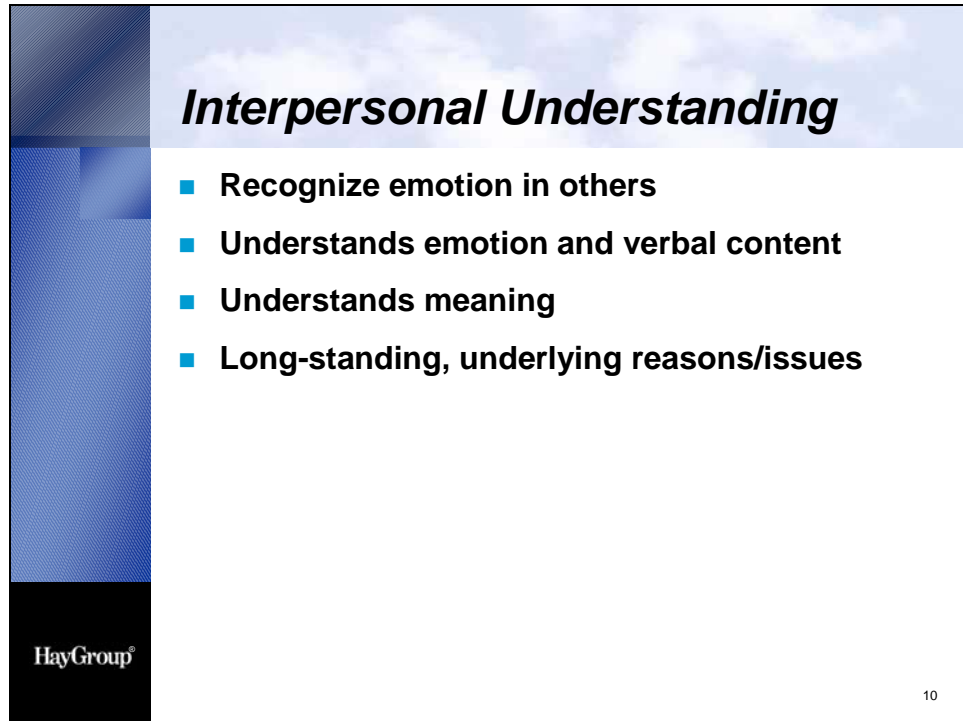
HayGroup®

9

Not to belabor, [but] let's look at one in a little bit more in detail, back to interpersonal understanding. In a minute I will spend some time about the lessons that we've learned that will apply to the emotional competence inventory. One of the things that we've learned is that there are various levels of a competency, increasing levels of sophistication. At the most fundamental level, we find recognizing emotions in others: people tend to do that, and they tend to comment on it, so, for example, you might read body language or tone of voice, or you'll notice when someone's upset or you'll notice when someone's anxious or nervous or sad, or whatever it is, it is essentially recognizing emotions in others. This competency is a lot harder than it seems, so various people have eluded to that. Essentially what it is, is understanding the person, really understanding the person, and when you get to the higher levels, the long term reasons and issues, that's when you really truly understand the person that you are interacting with, very difficult to do. So at the most basic level, we're sort of reading emotions in others.

I'm going to try to ground this in example of a story of a manager from a direct reports perspective. Direct report would interact with manager, and would notice lots of social anxiety, lots of awkwardness. When they had a conversation, the manager seemed to be nervous or awkward. It was relatively straightforward, but what the direct report also started to realize is that it made him uncomfortable as well. So that at the very most basic level of this competency, you might kind of call this competency "empathy," maybe.

Interpersonal Understanding

- **Recognize emotion in others**
- **Understands emotion and verbal content**
- **Understands meaning**
- **Long-standing, underlying reasons/issues**

HayGroup®

10

At the next level you understand emotional verbal content. Really what that could be is you understand partially expressed emotions or thoughts, [or] that somebody might say something to you and they're implying something but it's not clear what they're saying, and they may also be saying something else, and it sort of partially expressed, and it's not clear and the extent to which you can tune into the inconsistencies of what they're saying and what they're expressing (non verbally for example) gives you clues about what they really mean. So those kinds of things would be at the next level so, you know, the direct report would notice the manager would get anxious whenever the direct report wanted to talk about performance management--essentially whenever the direct report would say "hey, I want to set some goals," or "I'd like to think about what we can accomplish this year," or "what the next steps in my career in terms of my development?" and so forth. Having those kinds of conversations (sometimes not directly) would cause some sort of emotion in the interaction, because [of] some sort of anxiety, and the person would notice it and see this kind of pattern. So that is a connection that the person would make and that's at the second level.

At the next level you might understand the meaning behind that interaction. So: what does that interaction mean to the person? At this point the direct report [is] focusing on their perspective, but you really start to switch, and not so much focus on what you're feeling, or what's going on with you what it might mean to you, but really start to see it from the other persons perspective, and this is really where it gets hard. To truly understand where another person is coming from, it takes a lot of effort and energy, and you sort of have to suspend your own needs and, of course, there is all sorts of self evaluation going on--to what extent am I influencing this dynamic, to what extent am I perceiving accurately--and it gets to be a huge quagmire and at that point hopefully you are starting to really think about what is going on with this person, independent of what it means to you. Maybe you're getting triggered now because you have certain goals that you want to accomplish and this person is supporting you in another way. So maybe you might start to develop some insights based on interactions, and maybe you might determine… maybe the manager might have a general distrust, and he likes to have control of situations. Whatever the appraisal might be, it's going to be a way to help you organize all this information.

Finally, at the last level, longstanding underlying reasons or issues--that really is sort of understanding this interaction [in] context, so again, it's hard to say it any other way than I've already described, but [this is] really understanding where the person is coming from--maybe you might gather information from other situations, from other people, and you might start to think about the larger organization: you realize this manager has had similar difficulties and grievances over his career, and [you] actually might come to learn that what happened is that they brought in another person that took over this person's job, and the person wasn't effectively demoted, but they got a lot of their responsibilities taken away, which had a lot of their identity wrapped up into it. And so you start to piece together all this information and you get a bigger picture where this person is coming from and then you begin to appreciate more some of the emotional manifestations that you observed that might impact you as well. That is one example of what one of these competencies looks like.

This is a very rough outline, if we were to code it after reading an interview and a person happen to speak about something about would fall into this category the sort of the coding scheme would be much more complicated and sophisticated.

Okay, that is really the context. Lessons learned! So what have we learned from this that helps us understand how to access emotional competence and emotional intelligence competencies by this survey method?

One of the things that we've learned is that there seems to be scaling going on with these competencies, they seem to have an increasing levels of sophistication. I don't think that's anything new, I'm certainly not an expert in personality psychology, I don't know whole lot about personality measures, and I imagine that this has been observed, but I don't know the extent to which that this has been systematically assessed in the way that we've done. The competencies seem to be non-linear in the sense that the distance from one level to the next is not always the same. So making that first level observation-- noticing emotion--doesn't take a whole lot for the most part, although some people don't, and then getting to that second level… but the jump to that third level, that really takes something more, so that's what we mean by non-linear, non-cumulative.

What we've learned from studying competencies in this way is [that] sometimes within the same competency it's not more of the same thing. It's not always more of the same thing--like an influence: you make a rational argument to someone and you are trying to persuade them about something.

Lessons Learned
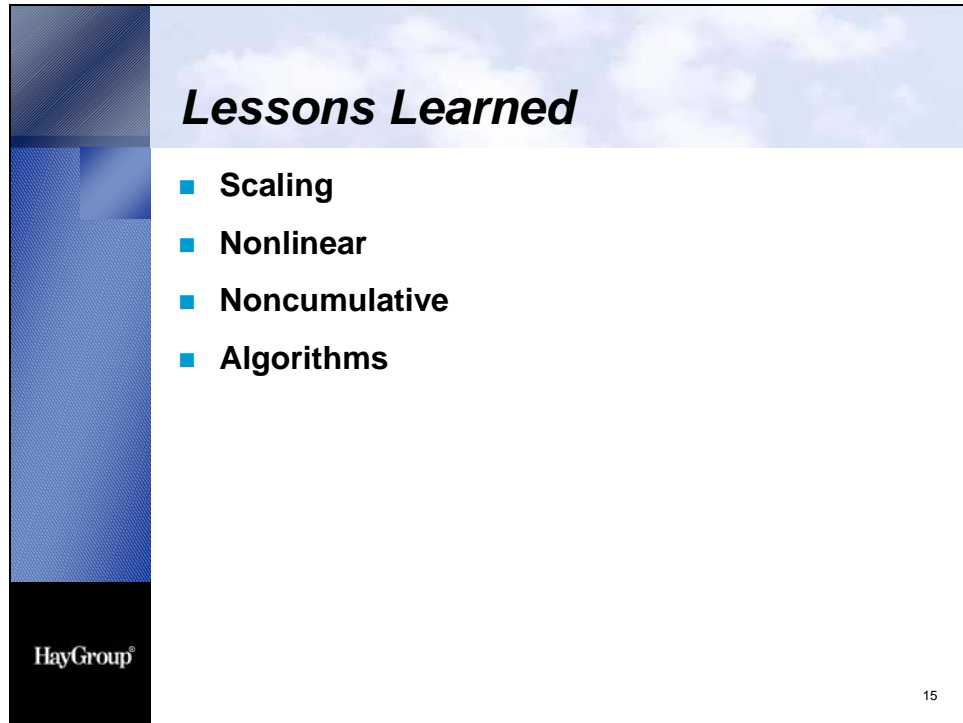
- **Scaling**
- **Nonlinear**
- **Noncumulative**
- **Algorithms**

HayGroup®

14

And we see people derail all the time, because what they do is that they will make this rational argument and they will be in a meeting with someone, it won't go anywhere, they will both leave and nothing gets implemented [and it's] wasted time, and it has not been an effective interaction. The person is banging their head and they go back and they get frustrated because they are not able to accomplish the things they need to get accomplished, and what they are doing is at an unsophisticated level so now they might say, "well, there's this and this but that doesn't work either," so what we find they are non cumulative sometimes it takes a very different behavior. It's still influencing but it looks very different, it's just not more of the same thing.

Another lesson learned is this idea of algorithms--this seems relatively obvious now that I'm reading it--you know there is more than one way to skin a cat. Because we really are in the business, we are really mostly interested in predicting performance; we work backwards, that's essentially what we do. We know what we want to achieve, because business has some sort of purpose, or some goal, whether it's literal, like, how do we increase sales, or whether it's more vague, like, how do we insure that we have good leadership in the future. There is a goal, then we work backwards, that's just how we approach it, obviously not a very basic science perspective.

So what we've seen is that people are successful doing very, very different things--the outcome is the same but obviously people are different. Personality dispositions are different--certainly we know that people achieve the very similar kinds of results in very different ways, using very different competencies, different behaviors, and different combinations of behaviors. One of the things we've learned is that if we're trying to predict performance, there really is no one formula, even within a particular organization within the culture--there are a lot of variability in terms of how people achieve those, so we have to be mindful of that when we're trying to help people get there.

So that's the context. Let's get into the ECI, so the emotional competence inventory. It's a 360 survey--what that means is we survey people, they do their self assessment--for the most part that's just for development reasons. Primarily again, we're looking to help people become more effective, let's say as a manager, so it's helpful for a person to assess themselves. And then we get ratings from managers, direct reports, peers, customers, others, and sometimes we get family members, but that's rare. We aim to get at least 9 other raters other than yourself, the more the better, 12, 15 the better. Essentially, the idea is they're a little more reliable, and one of the things we learned, like, for example, from the doctor-patient literature, when you are looking at patient satisfaction, which is a very important variable--you just don't want to ask one patient, you want to ask a number of patients, so you make sure you get varied view points. So 72 items, 18 competencies, not a whole lot of items, and unfortunately there are only 4 items per competency--we will see how that impacts reliability in a little bit.
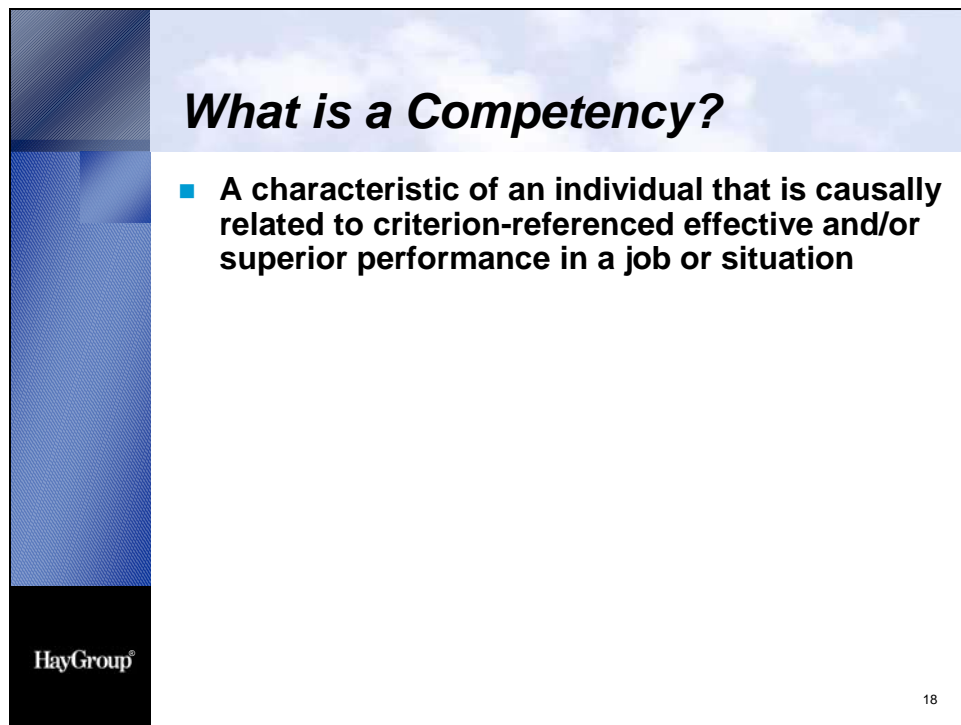
So the items really came from this interview method that I've spent some time talking about because we know what behaviors and skills and so forth seem to lead to outstanding performance at work; we sort of use that information to help us think about the items that might be most applicable to the content to what we're most interested in. So the items reflect the workplace context and we have 14,000 participants, and we've had two versions of the instrument we have about half and half. 7500 have taken the first version and now we have about 7500 that have taken this second version that seems to be a little bit better than the first-- I'm sort of thinking we need to revise again now and I'll explain why.

So the items are sort of behavioral frequency items; [each] describes a behavior and then asks how frequently do you see this person engaging in this behavior. So it might be something like "listen attentively:" do they consistently listen attentively, whereas maybe they don't listen attentively very frequently, so I think it's a five or six point scale.

What is a Competency?

■ A characteristic of an individual that is causally related to criterion-referenced effective and/or superior performance in a job or situation

HayGroup®

18

So let's just take a look at this again real quickly. Competency is characteristic of an individual that's causally related to criterion, reference, performance, superior performance in a job or situation, so that's really where we're coming from in terms of the competency.

What is an Emotional Intelligence Competence?

- **A competency that requires the recognition, understanding, or integration of emotional information about oneself or others**

HayGroup®

19

Now to make a jump to emotional intelligence competency: a competency that requires the recognition, understanding or integration of emotional information about oneself or others. So it's this basic idea of competency and thinking about how it requires emotional work of some kind, either understanding yourself, understanding others, or some combination.
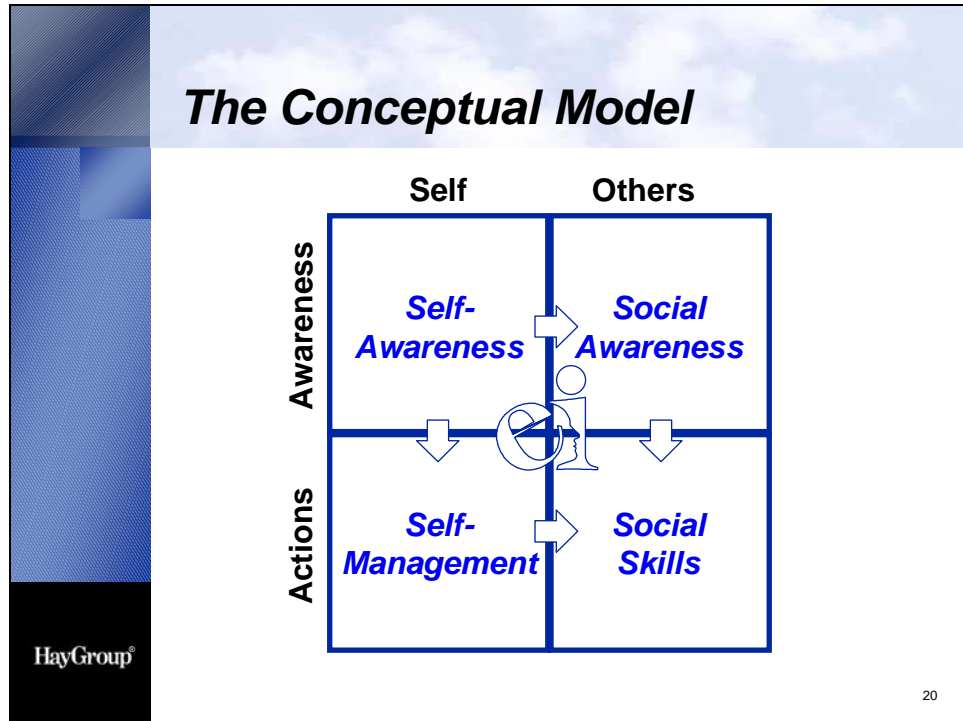
I wasn't going to show this but since Paul Ekman talked about the four characteristics of emotional intelligence--I'm not sure that's exactly how he phrased it, but it's nicely parallel—we can use this little model to help guide our understanding of what's going on here. In the upper left is a self awareness orientation which the idea of being more consciously aware of emotion when you become emotional. The social awareness piece is becoming more sensitive to how others are feeling; we sort of have this belief [that] we have one piece of research that supports this model, but we have this belief that being more consciously aware of your own emotions when you're becoming emotional helps you become more sensitive to how others are feeling. Then the lower left is self management--Paul said choosing how you behave when you are emotional, that is essentially what we're talking about The more that you have some ability in terms of self awareness you understand how you're more consciously aware of your own emotions--that's going to help you choose to behave under the conditions in which you are emotional. And finally social skills: using information about how others feel in a constructive way, that is sort of we believe driven by this model that we outlined here. This is the model that helps us understand what really is happening and how EI can help us make sense of workplace behavior.

ECI Sample Items

Competency	Item Text
Accurate Self Assessment	Acknowledges own strengths and weaknesses
Emotional Self Awareness	Knows how their feelings effect their actions
Self Confidence	Presents self in an assured manner
Achievement Orientation	Seeks ways to improve performance
Adaptability	Handles unexpected demands well
Emotional Self-Control	Gets impatient or shows frustration
Initiative	Initiates actions to create possibilities
Optimism	Stays positive despite setbacks
Transparency	Acts on own values even when there is a personal cost
Empathy	Listens attentively
Organizational Awareness	Understands the organization's unspoken rules
Service Orientation	Matches customer or client needs to services or products
Change Catalyst	Is reluctant to change or make changes
Conflict Management	Avoids conflicts
Developing Others	Recognizes specific strengths of others
Influence	Gets support from key people
Inspirational Leadership	Articulates a compelling vision
Teamwork and	Does not cooperate with others

HayGroup®

21

Here are just some items I figured why not just look at some of them specifically--I outlined a few in yellow, these are the ones that really get at what we think is the heart of what seems like is the heart of emotional intelligence, which seems to be consistent among various people that talk about it. So emotional self control, the third one down--gets impatient or shows frustrations--you know that happens all the time in the work place and it can lead to negative outcomes, it usually derails situations, sometimes it's an effective behavior, so it's not perfect--sometimes it's useful to get impatient or show frustration if you're trying to communicate a certain message. One of the things that is clear to us is that emotional intelligence clearly isn't being nice, or being good or being a pleasant person all the time; that is not the case at all. Certainly sometimes people that are highly emotional intelligent can be pretty harsh, actually. So that is one of the things that we've learned.

ECI Sample Items

Competency	Item Text
Accurate Self Assessment	Acknowledges own strengths and weaknesses
Emotional Self Awareness	Knows how their feelings effect their actions
Self Confidence	Presents self in an assured manner
Achievement Orientation	Seeks ways to improve performance
Adaptability	Handles unexpected demands well
Emotional Self-Control	Gets impatient or shows frustration
Initiative	Initiates actions to create possibilities
Optimism	Stays positive despite setbacks
Transparency	Acts on own values even when there is a personal cost
Empathy	Listens attentively
Organizational Awareness	Understands the organization's unspoken rules
Service Orientation	Matches customer or client needs to services or products
Change Catalyst	Is reluctant to change or make changes
Conflict Management	Avoids conflicts
Developing Others	Recognizes specific strengths of others
Influence	Gets support from key people
Inspirational Leadership	Articulates a compelling vision
Teamwork and	Does not cooperate with others

HayGroup®

22

I imagine you know we might want to talk about whether or not you can actually observe these things in other people, like do you know whether or not a person acknowledges their own strength and weaknesses. They have to sort of verbalize that, they have to have some sort of conversation around it. One of the things I should say, one of the two questions we ask up front or we believe that are necessary for this to be at all meaningful, is to you know you really have to be working with this person in a significant capacity--a few times a week you have to have what you believe is a significant interaction, and you have to report that you know this person well. And I know from working with various people that I can't really rate them on this because I just don't have enough interaction with them, so anybody that does not score high on those two items, we don't incorporate. I would love to talk about some of your thoughts on whether or we can actually observe that, like for example our emotional self awareness. You know how their feelings affect their actions, how do you know that? I think I can know that if I have enough opportunity to infer that, but it's not always clear. So those are some of the items.

Reliability Coefficients

ECI Cluster	Competency	Total Others Rating (N=6,595) Alpha Coefficient	Self Rating (N=6,354) Alpha Coefficient
Self-Awareness	Emotional Self-Awareness	.87	.71
	Accurate Self-Assessment	.82	.52
	Self-Confidence	.81	.71
Self-Management	Emotional Self-Control	.86	.71
	Transparency	.74	.52
	Optimism	.86	.68
	Adaptability	.81	.56
	Achievement Orientation	.80	.62
	Initiative	.70	.51
Social Awareness	Empathy	.89	.68
	Organizational Awareness	.84	.69
	Service Orientation	.81	.74
Relationship Management	Developing Others	.89	.73
	Inspirational Leadership	.90	.77
	Influence	.81	.63
	Change Catalyst	.83	.71
	Conflict Management	.54	.45
	Teamwork & Collaboration	.83	.56

HayGroup®

23

Let's look at some of the data, here are some reliabilities, this is the most recent version of the ECI, the 2.0. We made a bunch of revisions that seemed to help, and seemed to improve it a little bit, but after this latest round, I think we need to do a lot more and I'll push for that. Reliabilities are okay, the total others are all the peer and direct report and managers, those reliabilities seem to look pretty good, the self ratings have some problems--particularly conflict management. I think I know what's going on, I think there are a couple of items that are not clear, but I don't know. You have to keep in mind there are only 4 items, we have 72 items in the scale, we started with a 110 with the first version, we just have constraints, practical constraints, we can't have a 150 item survey, and have that fly--it just won't fly. So one of the things I'd like to do is cut down the number of competencies cause I think some of them are secondary, second order competencies. Factor analysis, it's a little shaky, there is basically one big factor that accounts for about 43% of the variance, so there is basically some global assessment of the person and what that is, we could probably disagree on, or debate. But it does a decent job of the 18 competencies if you try to confirm the 18 competencies, 14 of them 13 of them load pretty nicely, so there is something there, but clearly a difficult way. There is only a few measures that I've worked with that actually factor analyze so, anyway, [there is] more progress to be made on the instrument itself in terms of the psychometric.

ECI & Watson-Glaser (N=90)

ECI Cluster	Competency	Watson-Glaser
Self-Awareness	Emotional Self-Awareness	.04
	Accurate Self-Assessment	-.07
	Self-Confidence	-.13
Self-Management	Self-Control	-.13
	Trustworthiness	-.09
	Conscientiousness	-.15
	Adaptability	-.08
	Achievement Orientation	.05
	Initiative	-.12
Social Awareness	Empathy	-.08
	Organizational Awareness	-.07
	Service Orientation	-.20
Social Skills	Developing Others	-.23*
	Leadership	-.21*
	Influence	-.14
	Communication	-.13
	Change Catalyst	-.07
	Conflict Management	-.14
	Building Bonds	-.24*
	Teamwork & Collaboration	-.19

Murensky, C. L. (2000). The relationship between emotional intelligence, personality, critical thinking ability, and organizational leadership performance at upper levels of management. Dissertation: George Mason University.

HayGroup®

24

Here is some of the sort of construct validation stuff; I will talk about two studies. This is with the Watson-Glaser; you know the critical thinking ability, analytical reasoning test. Here is how I interpret it, it's open to interpretation, we wouldn't expect a strong relationship between these two variables. So for the most part, we are not seeing it, it's roughly around zero, there seems to be a sort of a negative pattern there, and this is a self report of the ECI's. So people rated themselves and then took this analytical reasoning test, and you can see sort of a little bit of a negative pattern and essentially not much relationship, so again, I think it's roughly consistent with what we expect and what we find, and I'll say this tentatively, in the lower quadrant there, the social skills, developing others and leadership and building bonds, are significantly negatively related: those that rated themselves higher on the ECI tended to have lower analytical reasoning skills on those. So one of the things that we do tend to see is that people that are more technically oriented have really strong analytical skills--and I know I'm essentially repeating the findings--but they tend to not be as good when it comes to managing people for example, we see people derail all the time that have technical skill and ability and have a background in terms of individual contributor background, if you know what I mean, and when they get in situations where they have to manage other people or be effective through other people, it's a huge shift, it's not a small shift, and a lot of people derail. It could be what's happening there, but I don't know.

Correlations with MBTI (N=18)

ECI Cluster	Competency	Extraversion/ Intraversion	Intuiting/ Sensing	Feeling/ Thinking	Perceiving/ Judging
Self-Awareness	Emotional Self-Awareness	-.26	-.62**	-.57*	.05
	Accurate Self-Assessment	-.44	-.53*	-.64**	-.14
	Self-Confidence	-.40	-.16	-.21	-.03
Self-Management	Self-Control	.01	-.57*	-.40	.06
	Trustworthiness	-.29	-.38	-.35	.10
	Conscientiousness	-.08	-.01	-.20	.36
	Adaptability	-.29	-.66**	-.53*	.07
	Achievement Orientation	-.16	-.22	-.19	.14
	Initiative	-.31	-.40	-.47*	.01
Social Awareness	Empathy	-.17	-.68**	-.65**	.09
	Organizational Awareness	-.14	-.37	-.44	.21
	Service Orientation	-.23	-.40	-.50*	.11
Social Skills	Developing Others	-.32	-.46	-.57*	.11
	Leadership	-.33	-.33	-.56*	.04
	Influence	-.20	-.41	-.48*	.06
	Communication	-.24	-.52*	-.53*	-.01
	Change Catalyst	-.32	-.39	-.50*	-.04
	Conflict Management	-.23	-.45	-.45*	.18
	Building Bonds	-.36	-.51*	-.60**	-.06
	Teamwork & Collaboration	-.30	-.60**	-.61**	.06

HayGroup®

25

Here are some correlations with the Myers-Briggs. So we're measuring personality probably to some extent, I think it does add a little value but to use our consulting language, but I will leave that decision ultimately up to you. So it seems like there is some overlap here, the top extraversion into intuitive feeling and perceiving--those are in a negative direction, so those that were rated higher on the ECI by peers managers and direct reports--four people minimum, up to more than 4 people rated these people--are senior level very high level executives from 18 paramedics. And so people that were high in the ECI tend to be more extroverted; it's not significant, you see sort of a negative pattern, you probably in the back can't read the numbers, but there decent effect sizes--relatively small but there seems to be something there, although not significant. People high on the ECI tend to a little bit more oriented toward intuition in perceiving information. Pretty strong relationship there; a pretty persistent pattern. In feeling, those high on the ECI used feeling more, used sort of the feeling orientation when making judgments used more personal and social values as opposed to logical reasoning, and not a whole lot with judging and perceiving. So there is clearly some overlap in terms of standard personality measures.

Self-Report: ECI & Big 5 (N=90)

ECI Cluster	Neuroticism	Extroversion	Openness	Agreeable ness	Conscientio usness
Self-Awareness	-.07	.47**	.28**	.00	.30**
Self-Management	-.20	.24*	.20	-.02	.33**
Social Awareness	-.10	.24*	.23*	.03	.21
Social Skills	-.11	.49**	.22*	.08	.39**

Murensky, C. L. (2000). The relationship between emotional intelligence, personality, critical thinking ability, and organizational leadership performance at upper levels of management. Dissertation: George Mason University.

HayGroup®

26

Another standard personality measures is the Big 5, this is the one with 90 executives from an oil corporation, 77 males and 13 females. Not a whole lot [of correlation] with neuroticism… these are both self-reports, so they rated themselves on ECI, and then rated themselves on the Big 5. Not a whole lot with neuroticism. There is certainly some overlap with extroversion, so people that are more extroverted tend to be higher on the ECI; a little bit with openness, a little bit with more openness, not a whole lot with agreeableness and consciousness. This wasn't the first version of the ECI, which had a consciousness subscale, which is no longer on the instrument, so that correlation might be reduced a little bit so there is clearly a lot of overlap, or a fair amount of overlap.

Bass Brewers Sales Agents (N=33)

Assessment	1	2	3	4	5	6	7	8	9	10
1. ECI Self Rating	--									
2. ECI Manager Rating	.59**	--								
3. Overall Performance	.62**	.51**	--							
4. No. New Dist. Pts.	.40*	.28	.85**	--						
5. No. New Accounts	.57**	.43*	.92**	.78**	--					
6. Customer Service	.04	.25	.38*	.35*	.16	--				
7. Annual Perf. Rating	.08	.23	.31◪	.40*	.15	.58**	--			
8. Years Bass Service	.09	.04	.04	.07	.07	.01	.26	--		
9. Career Progression	.39*	.23	.51**	.32◪	.33◪	.01	.10	.20	--	
10. Years of Sales Exp.	-.27	.07	.02	.04	.01	.04	.17	.73**	.13	--

Lloyd, M. (2001). Emotional intelligence and Bass Brewers Ltd. Dissertation: Nottingham Business School.

◪$p < .10$; *$p < .05$; ** $p < .01$

HayGroup®

27

Finally a couple studies about criterion validity: we looked at performance and a couple of organizations; let's start with this first one. This was Bass Brewers, really nice study, 33 people. These were area development managers. So within the entire U.K., these were the top 33, well, there is a smaller group, but these were the 33 area development managers, so they had different areas within the U.K. and they were responsible for generating business of course, for maintaining the relationships there, that they've already established, and looking to develop more business in the future. So these Bass Brewers sales agents, and the ECI ratings, we had both self rating and by their manager, so we just have one rater for the other category. And the nice thing is that there are solid performance metrics, it was designed to be sort of predictive, so they the assessments on the ECI happened six months prior to the performance numbers coming in and they are relatively objective. There is some subjectivity to them. The overall performance measure is all of the performance measures which I will explain in a moment all lumped together, so it's kind of the overall. The number of new distribution points that really is they measure the strength of the brand in that region, so you want to increase brand recognition within your region.

Call Center Agents: Correlations with Performance

ECI Cluster	ECI Ratings by Team Leader	Call Center Department	
		Client Services (n=44)	Sales (n=47)
Self-Awareness	Emotional Self-Awareness	.23	.33
	Accurate Self-Assessment	.38	.33
	Self-Confidence	.61*	.47*
Self-Management	Self-Control	.17	.26
	Trustworthiness	.66*	.53*
	Conscientiousness	.49*	.45*
	Adaptability	.37	.31
	Achievement Orientation	.64*	.35
	Initiative	.58*	.42*
Social Awareness	Empathy	.22	.42*
	Organizational Awareness	.49*	.25
	Service Orientation	.27	.39
Social Skills	Developing Others	.30	.30
	Leadership	.49*	.26
	Communication	.41	.32
	Influence	.53*	.37
	Change Catalyst	.57*	.43*
	Conflict Management	.45*	.26
	Building Bonds	.35	.48*
	Teamwork & Collaboration	.44	.41

HayGroup®

Nel, H. (2001). An Industrial Psychological investigation into the Relationship between Emotional Intelligence and Performance in the Call Centre Environment. Master's Thesis. University of Stellenbosch, Department of Industrial Psychology, South Africa.

28

So that's one. Number of new accounts! That's really sales, how many new distribution channels can you access, customer service, they had an outside organization do a survey of their customers to maintain good customer satisfaction, so that was the results, of the satisfaction from the customers perspective. Annual performance rating was done by the manager, so there is a little bit of a multiple common method bias here. And so that is based on mutually agreed-upon performance appraisal at the beginning of the year, and the beginning of the six months you sit down with your manager and say what do we want to accomplish by the end of these six months, you go back and forth, you negotiate, you agree upon "okay I will do this, this, and this, okay go ahead, touch back" and then at the end to the extent in which you achieve those goals, essentially like a five point rating scale. So there is some wiggle room there. Years of Bass service, career progression is the number of band changes, the number of promotions, divided by the number of years you've served; that is, your number of promotions -- then finally years of sales experience and years at Bass. A couple of interesting things strong correlation .59 and .62 under number 1, so [there's a] pretty decent relationship between the ratings on the ECI and overall performance.

Call Center Agents: Correlations with Performance

ECI Cluster	ECI Ratings by Team Leader	Call Center Department	
		Client Services (n=44)	Sales (n=47)
Self-Awareness	Emotional Self-Awareness	.23	.33
	Accurate Self-Assessment	.38	.33
	Self-Confidence	.61*	.47*
Self-Management	Self-Control	.17	.26
	Trustworthiness	.66*	.53*
	Conscientiousness	.49*	.45*
	Adaptability	.37	.31
	Achievement Orientation	.64*	.35
	Initiative	.58*	.42*
Social Awareness	Empathy	.22	.42*
	Organizational Awareness	.49*	.25
	Service Orientation	.27	.39
Social Skills	Developing Others	.30	.30
	Leadership	.49*	.26
	Communication	.41	.32
	Influence	.53*	.37
	Change Catalyst	.57*	.43*
	Conflict Management	.45*	.26
	Building Bonds	.35	.48*
	Teamwork & Collaboration	.44	.41

Nel, H. (2001). An Industrial Psychological investigation into the Relationship between Emotional Intelligence and Performance in the Call Centre Environment. Master's Thesis. University of Stellenbosch, Department of Industrial Psychology, South Africa.

HayGroup®

29

And then there's also a strong relationship between the self and other and that's unique--most of the time when we do something like this there is not a strong correlation between your self rating and your others rating or your manager rating, and there is not a strong correlation between a self rating and your performance. Managers do a little bit better job, but sometimes not. I think that one of the things we know about this organization is what is sort of called, a very high performance culture, they take performance very seriously, they sit down and they revisit their goals regularly, the managers and the sales people, and they meet very regularly to make sure that they are on track, that correlation although seems pretty high, makes some sense, it's nice to see that years of service and years of sales experience really don't relate to performance which sort of reaffirms this performance culture type of thing. So there are some decent correlations to performance in this setting. This is one with call center agents in South Africa and had really strong performance measures computer generated: your call service person and the phone call comes in and you have to manage that relationship, you need to give them information, you might cross sale or something so there is all kinds of… a handful of computer generated performance matrix for these people, and essentially what we see is some decent correlations between the performance measure and the ratings, and this is by their manager.

Call Center Agents: Correlations with Performance

ECI Cluster	ECI Ratings by Team Leader	Call Center Department	
		Client Services (n=44)	Sales (n=47)
Self-Awareness	Emotional Self-Awareness	.23	.33
	Accurate Self-Assessment	.38	.33
	Self-Confidence	.61*	.47*
Self-Management	Self-Control	.17	.26
	Trustworthiness	.66*	.53*
	Conscientiousness	.49*	.45*
	Adaptability	.37	.31
	Achievement Orientation	.64*	.35
	Initiative	.58*	.42*
Social Awareness	Empathy	.22	.42*
	Organizational Awareness	.49*	.25
	Service Orientation	.27	.39
Social Skills	Developing Others	.30	.30
	Leadership	.49*	.26
	Communication	.41	.32
	Influence	.53*	.37
	Change Catalyst	.57*	.43*
	Conflict Management	.45*	.26
	Building Bonds	.35	.48*
	Teamwork & Collaboration	.44	.41

Nel, H. (2001). An Industrial Psychological investigation into the Relationship between Emotional Intelligence and Performance in the Call Centre Environment. Master's Thesis. University of Stellenbosch, Department of Industrial Psychology, South Africa.

30

HayGroup®

We don't really hypothesis test--that's not really what we do; if we were to do some hypothesis testing, we know essentially what the requirements of that job are. Trustworthiness, for example, is key, credibility is key, if you are on the phone with someone and you feel like you can trust them and so forth you are more likely to be successful in that role. So we don't have the time, but it seems as though if we were to hypothesis test that it would be consistent with what we would expect, although we didn't literally do it, in this study.

Knowns & Unknowns

- **Knowns:**
  - **More Basic Research**
  - **Predicts workplace performance**
- **Unknowns:**
  - **What are we measuring?**
  - **Theoretical Foundation**
  - **Psychometrics & Statisticians**

HayGroup®

31

So I guess to conclude I guess we know a couple of things: we certainly would love to do more basic research, those that have expertise in this area, the unknowns piece, the psychometrics and statisticians we could use some help in running analyses and taking advantage of the data that we have and doing more basic research to really understand what's happening. We know that it is certainly predicts workplace performance, a couple of studies and we've seen it in a couple of other places, it predicts important moderating variables like organizational climate, which we know a lot about. Not really sure of what we're measuring, you know, whether or not we want to call it emotional intelligence that is sort of a convenient way that we can capture our understandings of competencies and how they work. There is not really a whole a really rigorous theoretical foundation, but it is much more empirically driven… we'd love to have a chance to think about that more carefully. So those are some of the things that we've learned; we certainly need to learn a little more, and one of the things that I'd love to do is the study with the MSCEIT and a real standard IQ test and some personality measures and some observational measures and ECI and of course performance ratings. I would love to do that and try to partial out the various relationships and I encourage collaboration with anybody that's interested. Thank you very much.

DISCUSSION

HayGroup®

DISCUSSION

DR. ROBERTS: Questions.

DR. KYLLONEN: Here is the question. You said that when you did the analysis, the first factor accounted for 43% of it …

DR. SALA: Yes, one big factor.

DR. KYLLONEN: And yet you have all these scales, and I'm wondering whether the user is the one responsible for requesting that kind of level of information, or whether the user is more interested in the single factor score?

DR. SALA: You mean when we apply it? When we get feedback?

DR. KYLLONEN: When you report back to the user, in other words the data say, that you are really only measuring one thing, but the user I guess is saying we spent all this money on your consulting company so we need about 25 things, is that what's going on?

DR. SALA: Well there's no question that it's helpful to them to think about these different kinds of behaviors, so they will get specific feedback on each of the competencies, and they will also look at them by each rater, so they'll see discrepancies between raters and who perceives what and how that differs from themselves, of course technically there maybe essentially one thing that we are measuring, but it is extremely helpful for people to get feedback that's hopefully candid from a variety of people that they work closely with, help them calibrate their own self perceptions. We are willing to kind of go with the error for the potential benefit to the individual, if that's what you're asking.

DISCUSSION

HayGroup®

DR. IRVINE: There's nothing like criteria in particularly if they are isomorphic; like the number of pints a beer drunken, which relates to the sales of the salesmen. It really worries about, a simple question, I assume that the correlations between these isomorphic hard criteria and the managers ratings is that correct?

DR. SALA: I'm sorry I didn't hear you?

DR. IRVINE: The correlations were with managers ratings, is that correct?

DR. SALA: Manager and self, yes.

DR. IRVINE: So that you are not experimentally independent because the manager knows exactly how much these people have sold.

DR. SALA: Of course, I said there is a common ...

DR. IRVINE: Indeed you've got what we would call autocorrelation among these variables. Do you tell your people that, do you tell your clients that?

DR. SALA: Do you tell them specifically about this study? We don't talk to them about this study; they are not necessarily interested. But to respond to that I, there is clearly a bias there, one thing I would love to do is to be able to control for the previous performance, because if it is predicted in terms of six months I assess them now and wait for the performance metrics to come in six months. As a manager, I have a sense of how successful they've been in the past. I know how effective they are, so that colors the influence in which I rate them. So, I think a nice improvement would be a control for previous performance and we haven't done that. We have done a handful, 3, 4, or 5 where the performance ratings for example which is very much subjective, very much your manager rates you on a 1 to 5 scale, almost like this annual performance rating that we have had here at Bass, and it doesn't correlate, so it doesn't always correlate.

DR. ROBERTS: Daniel Goleman, what's his role in this, because at some stage, he's listed as a consultant, he's working with this, he's not working with this, or what?

DISCUSSION

HayGroup®

DR. SALA: Dan Goleman has a long history with the HayGroup McClelland Center because he was I think it was in the 60's, he was one of Dave McClelland's students at Harvard, and so he knew David McClelland and had known Richard Boyzatis, who was once I think CEO or president of McBer, so they knew one another, and this work really came out of the work of the McClelland Center. But given Dan's ability to communicate in a way that is accessible to a wide variety of people that's how, I don't know if that answers your question, he got involved. He doesn't do the basic research, he doesn't sort of work directly with clients.

DR. ROBERTS: That's based a little bit on his theory of the actual dimensions of competence, or I think I heard that somewhere.

DR. SALA: Yes, so you mean his theory of competency, I would think that is largely determined from what he gleaned from the work that has been done there. But I can't say speak for that.

DR. ROBERTS: Anyone else?

DR. LEN WHITE: Maybe you can clarify this for me; McClelland, from my read, certainly preferred the projective technique to self-report, it is my understanding that there is some concern about the validity and value of self-report that led to the use of the behavioral event interview. What have you learned about the value of self-reports? Do they yield the same validity as say a behavioral event interview? Do you have any data like that that you can share?

DR. SALA: No, not specifically have we looked at that, I probably could actually get tons of data where people were assessed with a behavioral interview and then given some sort of additional assessment of them in a more sort of 360. But it might not work because there might be different competencies that were assessed and it's more sort of convenience. One of the things I think we feel pretty comfortable, or clear about, is that self-reports are valid. And because they certainly don't predict performance as we saw here, and tend not to correlate with how other people are seeing you (sometimes they do), we don't put a whole lot of stock in the self-report. Indeed, we primarily use it as a way to help people receive their feedback and calibrate their self-perceptions. Does that help? Does that answer your question?

DR. ROBERTS: Thanks Fabio.

PERFORMANCE-BASED EMOTIONAL INTELLIGENCE

DR. ROBERTS: I will now hand over the podium to Jack Mayer, who will talk about "Performance-Based Emotional Intelligence."

DR. MAYER: Good afternoon, I am Jack Mayer, from the University of New Hampshire. I am going to talk today about performance-based approaches to emotional intelligence.

I would first like to make some acknowledgements. In particular, I wish to thank Richard Roberts, Paul Gade, Pat Kyllonen, Gerald Matthews, Moshe Zeidner, Bill Strickland, and the folks at ETS, ARI, and HumRRO for inviting me. Between: I had a lot of fun explaining what these abbreviations stood for to people in my department, just before I left. I also want to thank my family for supporting me so that I might come and my collaborators on emotional intelligence: Peter Salovey and David Caruso.

My talk today consists of four parts.

Overview

- Theoretical Bases
- Model and Measurement
- Testing Hypotheses about Emotional Intelligence with the MSCEIT
- Conclusions

So I have four parts to this talk:

(1) Theoretical bases,

(2) The model and measurement we use for emotional intelligence,

(3) Hypothesis testing with the MSCEIT, and

(4) Then I'm actually going to spend a little time on conclusions and try to talk a little bit about what some other people said during the last two days.

Theory: Some of the Areas and Works From Which We Have Drawn

- Artificial Intelligence
  - E.g., Abelson, 1963; Dyer, 1983; Picard, 1997; Sloman & Croucher, 1981
- Clinical Psychology
  - E.g., Sifneos, 1975; Taylor, Ryan, & Bagby, 1985
- Cognition and Affect
  - E.g., Alloy & Abramson 1979; Averill, 1983; Bower, 1981; Isen, Clark, Shalker & Karp; 1988; Schwarz, 1990; Zajonc, 1980
- Emotion, Appraisal, and Evolution
  - E.g., Darwin, 1872/1998; Davitz, 1969; Ekman, 1973, 2003; Plutchik, 1984; Roseman, 1984; Scherer, 1993

- Intelligence Testing
  - E.g., Detterman, 1983; Gardner, 1983; O'Sullivan, Guilford, & deMille, 1965; Sternberg, 1985;
- Neuropsychology
  - E.g., Lane, et al., 1990 TenHouten, Hoppe, Bogen, & Walter, 1985
- Nonverbal Communicatio
  - E.g., Arnheim, 1974; Buck, 1984; Nowicki & Mitchell, 1998; Rosenhan & Messick, 1966
- Philosophy
  - E.g., Calhoun & Solomon, 1984
- Contemporary Commentators
  - Barrett, Izard, 2001; Roberts, Zeidner, & Mathews, 2003

Unfortunately, though I would love to, I can't spend too much time going through all the amazing contributions that people, both in and outside this audience, have made to our original work (and continue to make to our thinking about emotional intelligence). However, I did want to talk about some of the major areas.

When we wrote our 1990 articles (one on measurement, the other theory) we were drawing together disparate literatures. These included research from the artificial intelligence literature (articles like why should robots have emotions); clinical psychology work in alexythemia; Richard Lane's Levels of Emotional Awareness; and studies at the interface of cognition and affect (which I certainly was trained in a number of years, Peter as well). We also looked at the research on emotional appraisal and evolution (Klaus Scherer, Paul Ekman and a number of others); intelligence testing (Robert Sternberg, Doug Detterman, O'Sullivan and Guilford, Maureen O'Sullivan spoke to you earlier today); neuropsychology (Richard Lane, TenHouten, and Walter); non verbal communication, and those people who continue to inform our work as contemporary commentators, including Lisa Feldman Barrett, Richard Roberts, Gerald Matthews, and so forth.

Emotion as Information

- Emotion as an "object" of intelligence
- Emotions convey information about relationships
- Emotions can also operate as an auxiliary storage and communication system parallel to intelligence

Sources: Mayer & Mitchell, 1998; Spearman, 1927; Schwarz, 1990; Zajonc, 1980

My introduction has already given a lot of the theoretical ideas and that's okay. And I am going to go through them a bit more quickly than I might otherwise, partly because I have more information about the emotionally intelligent individual that I would like to share with you, as identified by our test, the MSCEIT.

But we agree with Paul Ekman, Klaus Scherer and many others that emotion conveys information and our view [is that] emotion is the object of intelligence, but also emotion can facilitate intelligence, which I will talk about in a little while. The information that emotions convey is information about relationships. Emotions can also operate as an auxiliary storage and communication system, parallel to cognition--these are ideas that a number of emotion researchers have put forth--we don't claim them to be unique to our research team.

Emotions are evolved signals about relationships (and, hence, universal)
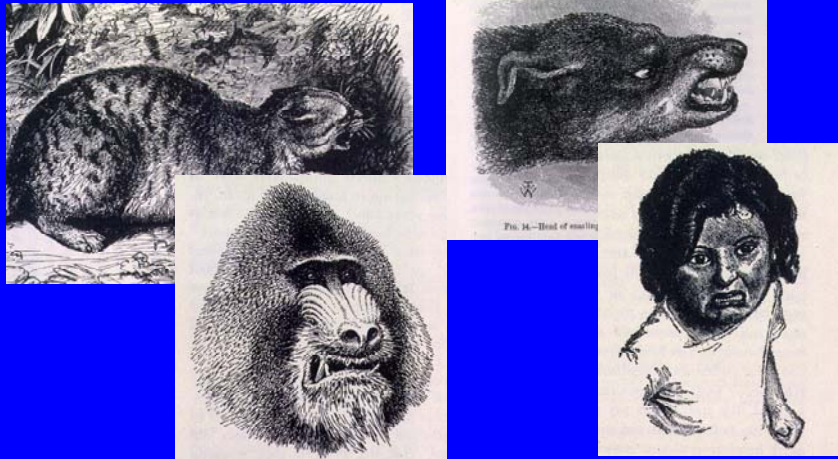
Sources: Darwin, 1872/1998; Ekman, 1973

I thought Paul Ekman was going to talk about this today, but I'll mention it briefly. Emotions are evolved. Darwin came up with the idea but Paul was really responsible for proving Darwin's ideas and demonstrating the generality of Darwin's ideas. And another way of thinking about emotions and the information that they convey is through appraisal theory, which Klaus Scherer talked about earlier.

Specific Emotions Communicate Specific Information

If the emotion is:	...then the individual's appraisal involves:
Happiness	Security; pleasure in joining others
Sadness	Loss
Fear	Being threatened
Anger	Injustice

Sources: e.g., Ortony, Clore & Collins, 1988; Plutchik, 1994; Roseman, 1984; Smith & Elsworth, 1988; Scherer, 1993.

Each emotion has a set of sub-emotions, and each sub-emotion is associated with a particular appraisal. This is extremely elementary, Klaus already got into more detail than I have here, but the idea here is fairy simple. Take happiness as an example. If you appraise people securely, if you have pleasure joining in others, that is going to be associated with happiness. Sadness, by comparison, is associated with senses of loss; fear is associated with threat; while anger relates to injustice.

I am going through the intelligence material because there are lots of different possible perspectives on intelligence, and it may help you to have some sense of where we're coming from.

I don't know whether this to say we or I here, so I will say the later. I believe that *g* relates to the capacity to carry out abstract reasoning, it involves knowing the similarities and differences among concepts, the capacity to make generalizations about concepts, and I believe that *g* provides a very good description of a general unitary dimension that describes a variety of cognitive area types of reasoning. In addition, I believe that you can model intelligences so it makes sense to talk about certain specific areas of intelligence: verbal intelligence, perceptual-organizational intelligence, spatial intelligence, and yes even, emotional intelligence.

<div style="background-color:blue; color:white; padding:20px;">

The Scope of Emotional Intelligence Involves:

- The capacity to reason with and about emotional signals
- The capacity of emotion to enhance thought

</div>

So if we are going to draw an analogy--to pick up on what one of the earlier speakers, said--if we're going to draw an analogy from verbal intelligence, or spatial intelligence or perceptual-organizational intelligence … if we are going to draw an analogy from that to emotional intelligence, then emotional intelligence ought to be one of two things, or both. It ought to be the capacity to reason with and about emotional signals, about emotional information, or, secondly, the capacity of emotion to enhance thought.

Now, I think actually we can probably get pretty good agreement on that statement then we can move into, well okay, if you are going to do that, what are some of the specific areas you put under this emotional intelligence idea? And there we may have some disagreements; I think there is probably more than one valid way to talk about the specific areas. Klaus suggested some ways, Paul Ekman suggested some ways, but we'll come back to that in just a moment.

> ## A Digression Concerning the Popularization of EI
>
> - In 1995, Daniel Goleman popularized EI in a best-selling book
> - The book subtly but importantly changed EI's definition.
> - It also made claims about EI's predictive power that my colleagues and I had never made
> - Some researchers and critics employ the popular definition without realizing it is different …
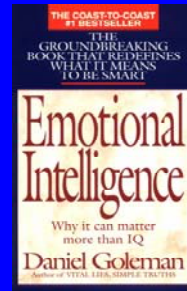
I do want to mention briefly the popularization of emotional intelligence. In 1995 Dan Goldman of course published his best selling book *Emotional Intelligence*, and ostensibly the book was based on our 1990 theory.

As I've written in Bob Sternberg's *Handbook of Human Intelligence* we changed our definition a bit from 1990 to 1997.

Dan also changed our definition of emotional intelligence but in the opposite direction. I would like to think we made ours clearer, more focused, and [we gave it] more heightened utility. I think that Dan changed the definition in another positive way--to communicate well, communicate generally by a whole host of different researchers. Dan's book on emotional intelligence started out, although it essentially was representing our 1990 definition, as a very broad representation of it indeed, including the fact that emotional intelligence was character, which is something we never said.

It also made claims about EI's predictive powers that my colleagues and I had never made. Since then this area has come up very, very quickly. Sometimes we say we've actually constructed a whole new area of intelligence studies in about 15 years. In a way, we've gone through the whole progress of a century's worth of intelligence research, and we wanted to suddenly expect emotional intelligence research to catch up.

Models and Measures: Four Branch Model of Emotional Intelligence
(Mayer & Salovey, 1997; Salovey & Mayer, 1990)

Emotional Intelligence

- Managing Emotions
- Understanding Emotions
- Facilitating Thought
- Perceiving Emotions
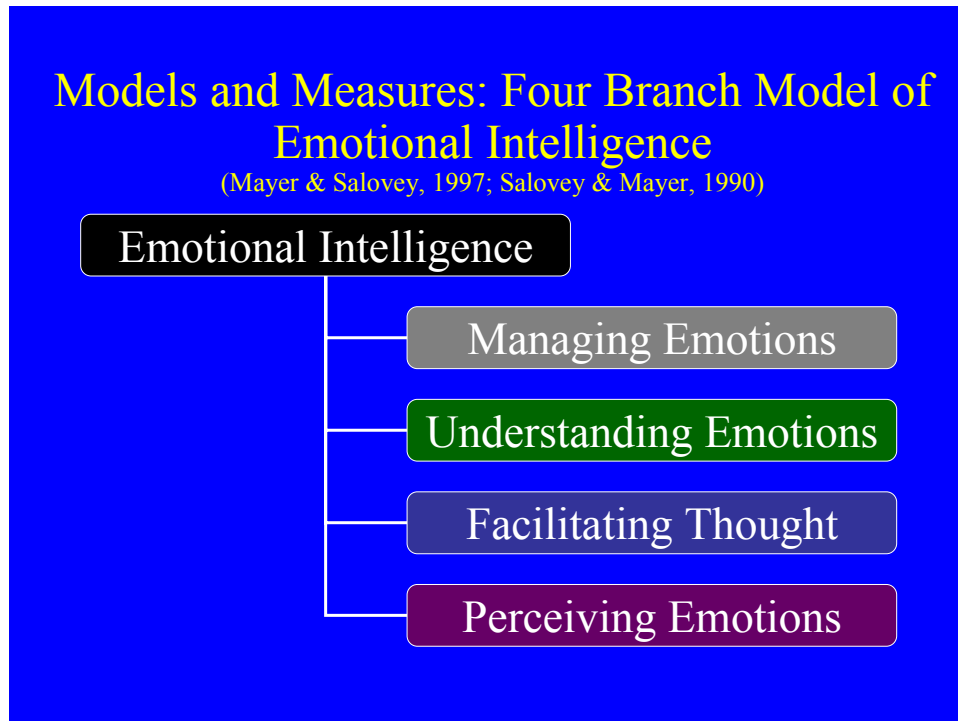
So what is the model of the EI as we see it, and what's the way that we measure it? Again, I think that there are different ways to carve up the world, there's not just one way. But in our reading in 1990 of the emotions research, non-verbal communication research by Paul Ekman, Klaus Scherer and others, and research more broadly in the clinical literature and elsewhere, we decided that there were three areas that probably could be pulled together into a potential domain of emotional intelligence. Then, in 1997 we made it four areas.

And these are the four areas, and I recommend this current definition of emotional intelligence over the 1990 one, or along with the 1990 one. We say that emotional intelligence involves four areas: perceiving emotions, using emotions so as to facilitate thought, understanding emotions, and managing emotions.

These Four Branches Can Be Illustrated with the MSCEIT Test…

- To take the test, number from 1 to 6 on a page.
- Record your answers to the following questions…
- Note: This is only an illustration; it is too short to measure emotional intelligence with confidence (technically, it lacks high reliability).
- Its purpose is to illustrate how the longer tests work.

And I'll talk to you a little bit about each of those four branches, and as I do, I'm going to illustrate them with some test items. I apologize to about a quarter of the audience who I can tell just by facial, visual recognition have already taken this test. But for the benefit of portions of the audience, who haven't had that chance, take with me some items of the MSCEIT, as I talk about these four areas of emotional intelligence (I think the other members of the audience will be patient as we do that). So to take the test, number from 1 to 6 on a page, if you are interested… this is what probably will be the most sophisticated psychometric audience I could possibly speak to, but I'll just mention with just 6 items you are not going to get a reliable score, our test actually has 141 items to yield a reliable score.
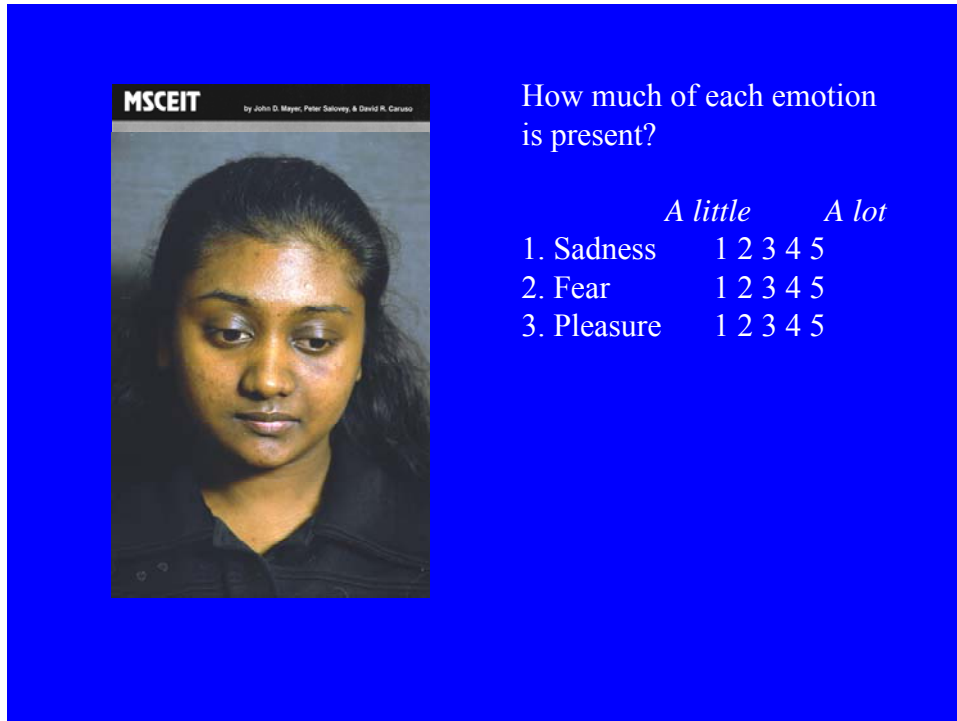
The first idea is that emotional intelligence involves the identification or perception of emotion and about the only thing that I can contribute--and it's a very minor contribution on top of things that have already been said about that--is that the development of emotional identification we have this wonderful thing called empathic mirroring where parents reflect the facial expressions of their infants, or mirror some of the emotional expressions infants make.

They don't mirror it all the way; here is sort of a frantic child, the mother herself is not frantic, but maybe she is going oh, oh, oh, calm down a little bit.

How much of each emotion is present?

	A little			A lot
1. Sadness	1 2 3 4 5			
2. Fear	1 2 3 4 5			
3. Pleasure	1 2 3 4 5			

How much of each emotion is present in this face? Sadness from a little to a lot, if you're doing this along with me, answer how much sadness is in that face, one is a little and five is a lot. That's item one. Item two is the same thing, but do it for fear. How much fear is in that face, 1-5, five for a lot of fear. Item 3, how much pleasure is in that face from 1, none at all, to 5, a lot. And that's one way that we have of measuring emotional identification in the MSCEIT. We will score [this] in a little while.

Branch 2: Emotional Facilitation of Thought
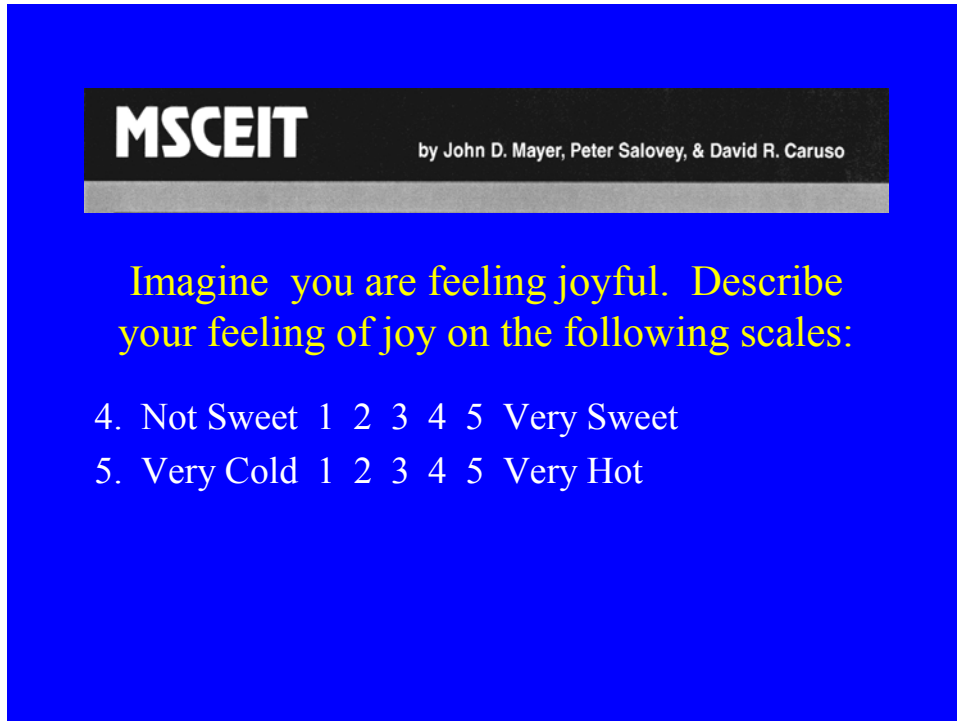
- Emotions Prioritize Thinking
- Emotions are Available as Aids to Judgment
- Mood Swings Change Perspective

ILLUSTRATION: MUNCH'S "THE HOWL"

The second branch of emotional intelligence is using emotions to facilitate thought. Sometimes our model is called a cognitive model, maybe that's right, speakers use the expression … I'll leave it to you, I guess, to decide. That kind of terminology sort of rubs me the wrong way--I don't think of our model as a cognitive model, I think it as a joint emotional-cognitive model. The same way that having, for example, a large vocabulary evokes a sign of high verbal intelligence. It's the same thing with emotions; the more you know about emotions, the more you can use … your emotional knowledge to help your thinking.

So for example, if you know that when you're sad you're more analytical, and maybe a little bit less motivated, that is maybe a good time to do writing, because writing is an analytical task. It makes you a little bit less motivated to get out of your chair and do something else, [so] it might help you stay and write a little bit longer. So emotions can facilitate thoughts if you know about them. Emotions also prioritize thinking, and they can be available as an aid to judgment. In this particular painting, The Howl, I'm told that Munch actually felt pretty bad when he painted this and he used this idea of emotions as a parallel system, drawing on that information from his emotion system as he was painting.

MSCEIT by John D. Mayer, Peter Salovey, & David R. Caruso

Imagine you are feeling joyful. Describe your feeling of joy on the following scales:

4. Not Sweet 1 2 3 4 5 Very Sweet
5. Very Cold 1 2 3 4 5 Very Hot

We measure this in a kind of interesting way; I'll show this to you just because I think of this as a kind of interesting task. It is called the synesthesia task, and we ask you to compare emotion sensations to other sensations… for example, imagine you are feeling joyful, describe your feeling of joy, as to how sweet it is. And by sweet we mean taste sweet, how sweet does it taste from 1 not sweet at all to 5 very sweet. This is item 4. And for item 5, imagine you were feeling joyful and here we actually ask the people to create a small sense of joy within themselves and them compare it to how cold or hot you feel, one would be very cold, 5 would be very hot. Anyone need more time?
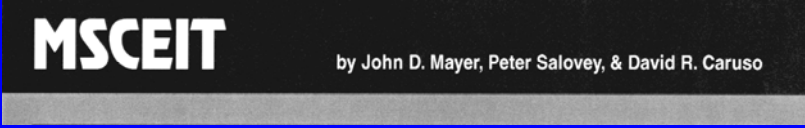
We're actually understanding and analyzing emotion. This is another place were there is a tremendous amount of emotions research and appraisal research [that] really relates to this, but there also some stuff which is slightly disembodied from appraisal research which relates to emotional understanding. This is one of my favorite diagrams by Robert Plutchik, an early evolutionary emotions theorist, in which he arranged emotions in sort of emotions circle analogous to a color circle. I just thought this was a kind of innovative thing for him to do. The vertical axis […] represents intensity, so I think he was using a fair amount of emotional understanding to create this diagram.

Using that particular diagram as an expert criterion source, we ask the question, what emotions might combine together--because this is another part of Plutchik's theory--what emotions might combine together to make the emotion of contempt. I have to be careful when I say this in front of emotion researchers, because some acknowledge that there are blends and some don't. If you will allow me that leeway I'll take it. So here you answer a, b, c and d. What emotions might combine together to make contempt: a, anger and fear, b, fear and annoyance, c, anger and disgust or d, disgust and joy?
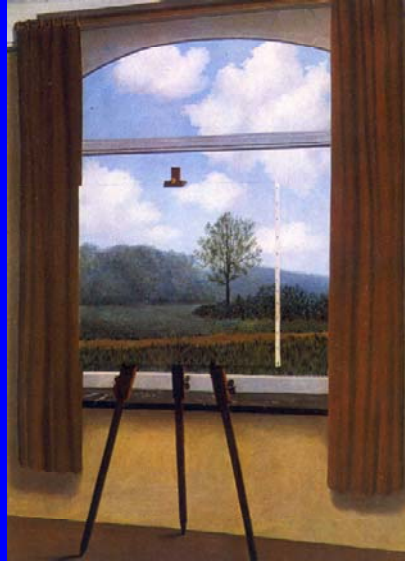
Branch 4 involves regulating emotion. One of the concerns I have with the popularization is the emphasis that is placed on emotional control. Now emotional control is important, as you know when you're raising a child it's important to have the child be able to control his or her emotions to some extent, but there is also such a thing as over-control, and it seems like the popular emphasis, the media emphasis was on how does an executive really control her emotions at work, or really control his emotions at work, and while you do have to do that, no question, you also have your let your control loosen up so you can understand what happens when you do express your emotions, or how to express your emotions and so forth.

I don't have an item for this because it's too long--those items are a little bit longer, those items actually get closest to the situational judgment tasks that I think ETS is interested in.
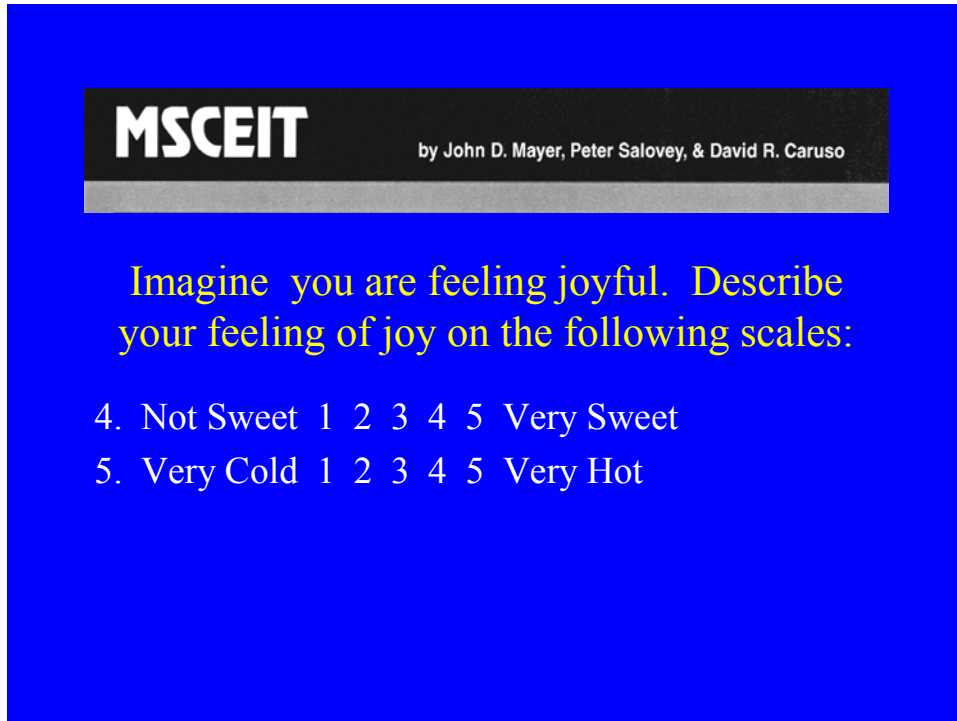
So now I have a not so hidden motive, in asking you to take these six items. I want to demonstrate one form that we have of identifying a criterion for the correct answers to these items, which becomes an issue.

Let me ask you, those of who took this test just now, how many people for sadness, said there was a little sadness in this face, number 1, raise your hand? (I think that Doctor Ekman had a meeting, so he had to leave the room, so you don't have to worry about him judging you). How about 2, raise your hand if you put 2, how about 3, a few courageous souls, thank you very much, about 5 or 6 people, how about 4, okay, that's a lot, how about 5, a little bit less than a lot. Okay, so here, it being 4, we call this consensus scoring we give you credit for 4 but we also give you a little credit for 5 cause we're nice people and so if you put a 4 or 5 you get credit for that item, how about fear? How many people put 1 for fear? A medium amount, how about 2? Also medium, 3, far fewer, so let's cut this a little quick and we'll go to 1 or 2 for item 2. How about pleasure, how many people said 1 for pleasure, that's a medium amount, 2 for pleasure, a little less, 3 for pleasure, still 3, 4 for pleasure, 5 for pleasure… right about here we are no longer measuring emotional intelligence, we're measuring social courage!

Imagine feeling joyful, how many people said 1 for not sweet for joyful, 2, 3, 4, medium and 5 medium, so the answer here would be 4 or 5. Item 5, how many people said joy was very cold 1, 2, 3, okay, 4, this is sort of medium, this is a little bit larger than medium and 5 a little bit dropping off, so three or four for how hot joyful is. Klaus, you did this all across Europe, something like what we're doing right now, correct?

DR. SCHERER: Yes.

DR MAYER: Yes, did you, am I remembering properly that [on] one of those large European surveys, you sort of collected some data like this?

DR SCHERER: Yes.

DR. MAYER: And then finally what emotions might combine together to make contempt, a. anger and fear, b. fear and annoyance, c. anger and disgust, or d. disgust and joy. What would you say?

AUDIENCE: C.

DR MAYER: And so, in most places people answer, "C", I guess it's partly a function of our great translators, but we know that in Japan, in Croatia, in Israel, In England, in Norway, people generally answer C to this item. […] In Saskatchewan, Canada, very interestingly, there is a sizeable and vocal minority who believe that the answer is "A" in Saskatchewan, and that is the only place I've ever been where I've seen that cultural difference. But there are certainly cultural differences we will talk about that as we go along. So this is to give you some idea, this is one way we score the MSCEIT with consensus scoring.

Peter Legree is going to talk about consensus scoring later today so I'm not going to belabor it. But I want to say this is actually our preferred way of scoring the test. But other people have another way that they prefer which is with expert scoring, and we've also investigated that alternative.

How Was the Test Standardized?

- Standardized on 5000 Participants Across over 50 English-speaking data sites in:
  - Australia
  - Canada
  - India
  - South Africa
  - United Kingdom
  - United States

- Ages 17 to 79
- Reports matched to United States Census Data on age, gender, ethnicity and education

So the MSCEIT now has been standardized. The original standardization sample is about 5000 people across 50 English-speaking data sites in a number of different countries. We have ages 17-79, and some of the scales have been matched to the United States census data.

TESTING HYPOTHESES

Q: Do General Consensus and Expert Evaluation Converge On the Right Answer?

The prediction:
- Experts and general consensus should often converge because of the evolutionary-meaning hypothesis.
- If they don't converge, at least in part, it throws "correct answers" into doubt

Sources: Legree, 1995; Roberts, Zeidner & Matthews, 2001; Mayer, Salovey, Caruso, & Sitarenios, 2001

My, how can I describe… my calm equitable colleague, Frank Landy, in a not so recent paper, a paper around 1980 suggested that we no longer talk about tests' reliability and validity, but rather we just acknowledge that we're doing hypothesis testing with our tests. […] We claim that this is content valid, that is to say that each of the four branches, measured each with two tasks, [is content valid]. We had the test called the MEIS in the late nineties, where we had 12 tasks, which was probably better for factor analysis and stuff like that, but we also wanted to be practical and a twelve-branch 400 hundred item performance scale is not very appealing to people, so this is much more expedient. And so we want to ask questions about both the tests and about emotional intelligence like does emotional intelligence exist? Does the test work and so forth?

And in the spirit of Doctor Landy, I'm just going to talk about some of the hypotheses we've tested as we've gone forward with the MSCEIT. First question is: How does the consensus scoring that I just went through with you compare in some way with what results we would get if we went out to experts and asked emotions experts what the correct answer was to MSCEIT? And in fact we asked 20 members of the International Society for Research in Emotions--of which I bet we have a number of members here!--to please take the MSCEIT and show us what the correct answers were.

Roberts, Zeidner, and Matthews and others suggested if those criteria do not converge, we're in trouble. So our first question is: Can you get a right answer? Basically I think [what] this boils down to is, is there a right answer to these test items?

A: Yes, There is Convergence Between General and Expert Scoring as to the Right Answer

- The MSCEIT *r for agreement* between group consensus and experts ranges from *r = .90* upward
- Where Experts Beg to Differ
  - Largest differences from general group in Branches 1 and 3 (e.g., faces and understanding)
  - This is where there is greater understanding of emotion
  - Emotion theories are best worked out in these areas

Source: Mayer et al. (2003), Emotion.

And the answer is yes, there is convergence, the convergence between group consensus and experts ranges .90 and upwards. We didn't want it to be 1! We want experts to know something that the group doesn't know. There are actually structural differences between experts and [the] general consensus. Experts agree with one another a lot more.

And another small difference between experts and the group is the experts are particularly better on branch one, where certain people who were in this room just a few minutes ago have us instructed us well in how to do those tasks, and they are also better on branch three, where other people in this room have instructed so well about what emotions mean.

Experts are great on branches one and three where the most research is, they are a little bit less good on two and four where there is the least--at the time that we did this, there was less research.

Q: Is the Test Reliable?

Second question: is the test reliable? One of the things we were very attentive to in looking at the early nonverbal perception tests coming out from a number of laboratories as reviewed by Buck in 1980 (or so) was that they suffered from issues of reliability.
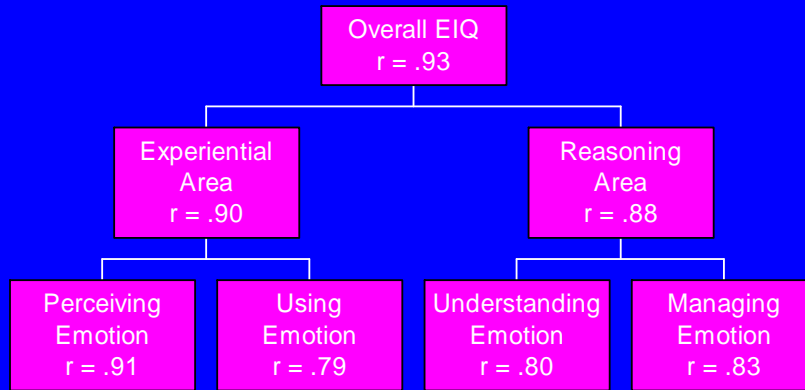
And I won't go into the long song and dance about how we fix that, but we did fix that and so when we look at the MSCEIT for example we see that the overall test reliability is .93 [in] this study. Then in the branch level we have reliabilities of .91, .79, .80 and .83.

Now lot of people say they've heard [this test] is not reliable and where is that coming from; well, remember, there is a lower level of tasks, two tasks for each branch as you can tell. The tasks are less than optimal for individual psychodiagnosis or psychological assessments--we don't advise you to interpret the tests at the task level, we advise you to interpret it at the branch area, or the total test level.

Another question, is EI a unitary intelligence? That is, do various EI tasks load on a single EI structure? Now here I've got 8 tasks to the MSCEIT-- actually in the late 90's, we were using the MEIS, which had 12 tasks--I'll just let you [know] that the results here are the same as we found for the MEIS in 1991.
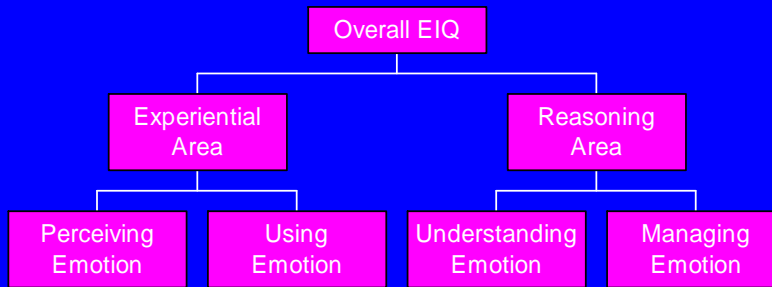
All this really says is yes, you can successfully model these 8 paths as a unitary ability. Now fortunately, here in 2003, I would like to think we are a lot more sophisticated than we were in 1980 or 1990 when we were having cut-throat arguments about is intelligence one thing, or is it several things? These are mathematical models and you can do it both ways, so we can model it as one unitary validity and that works very nicely, and we get improved fit if we model it alternatively as four highly correlated factors.
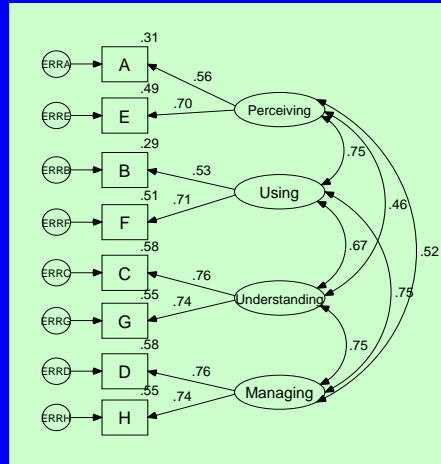
The third question is: Is there support for the four branch model of EI from confirmatory factor analysis?

A: Yes, the 4 Factor Model Shows a Good Fit

Goodness of Fit Index (GFI)	0.989	Adjusted GFI (AGFI)	0.972
Normed fit index (NFI)	0.977	Relative fit index (RFI)	0.958
Tucker-Lewis index (TLI)	0.964		
RMSEA lower bound	0.042	RMSEA upper bound:	0.062

Source: Mayer, et al. (2003) Emotion

And the answer is, to continue with what I was saying, yes you can do it both ways, if you look at these branches as four intercorrelated areas of ability, you get also a very nice fit, in fact you get somewhat superior fit. You can do your chi-square and there is a significant improvement. I'm not showing you the chi-square, because you know we got 5000… actually I think this is on the first 2000 on those first 5000 subjects… when you have 2000 subjects the chi-square will always be significant because it's sensitive to N.

So it can be scored with correct and incorrect answers. It is reliable--you can talk about a general emotional intelligence, you can divide it into specific related areas.

Q: Does the MSCEIT Measure Something New Relative to Earlier Tests?

•Tests that measure similar or identical concepts typically correlate in the r = .65 to 1.00 range.

•We want the MSCEIT to measure something new – to be independent of earlier scales – to correlate in the r = .00 to .35 range.

Does the MSCEIT measure something new, relative to earlier tests? Tests that measure similar or identical concepts typically correlate to the .65 to 1.0 range, we want the MSCEIT to measure something new, to be independent of these earlier tests… is that the case?

A: Apparently So. the MSCEIT is Distinct from all the Scales Below (N's > 100):

☯ Intelligence Tests	*r* = .00 to .30
☯ Big Five Personality Scales	*r* = .00 to .35
☯ Self-report Scales of EQ, optimism, empathy	*r* = .00 to .35
Sources: Ciarrochi, Chan & Caputo, 2000; Mayer, Caruso, & Salovey, 1999; Roberts, Zeidner, & Mathews, 2001; Salovey, Mayer, Caruso,& Lopez, in press.	

And the answer is yes. You know those yin yang things, I didn't put those here, I had check marks, so somehow from my Window Power Point 97 to this it has become yin yang. I am going to be high on Gerard Saucier's dimension of spirituality for the next few minutes.

AUDIENCE: [laughter]

DR. MAYER: With intelligence tests, correlations are about .0 to .37 or so, .37 specifically between our branch 3, understanding emotions, and verbal intelligence. Zero between branch 4 and Ravens Matrices. Kim Barchard, who is now an assistant professor at the University of Nevada at Reno, used the O'Sullivan and Guilford scale to look at its relationship to the MSCEIT, and shockingly--I have to qualify this, she used only pieces of the O'Sullivan and Guilford and only pieces of the MSCEIT. But apparently the relationship is not all that high, her correlation was at the .20 level, .19 level so even correcting for attenuation this seems to be different than social intelligence.

A: Apparently So. the MSCEIT is Distinct from all the Scales Below (N's > 100):

☯ Intelligence Tests	*r* = .00 to .30
☯ Big Five Personality Scales	*r* = .00 to .35
☯ Self-report Scales of EQ, optimism, empathy	*r* = .00 to .35
Sources: Ciarrochi, Chan & Caputo, 2000; Mayer, Caruso, & Salovey, 1999; Roberts, Zeidner, & Mathews, 2001; Salovey, Mayer, Caruso,& Lopez, in press.	

We've got rumors that [this is] different than practical intelligence too, but just rumors, I think. It is much different than the Big 5 personality scales, which shouldn't surprise you—right, that's self-report, [whereas] this is an ability measure, all that kind of good stuff. It's also utterly different from self-report scales of emotional intelligence, or EQ as it is popularly called. Right after the popularization, there was a whole rash of self reported emotional scales that came out and we found that those have correlations with the MSCEIT between .35 (actually a traditional emotionally oriented empathy scale) to down to .20 down to .0 depending upon the scale.

Actually somebody wrote a self-report emotional intelligence scale that has negative correlation with the MSCEIT. So it's really independent of a large number of scales that have been developed in the past. And that made us both happy and nervous!

Q: Does the MSCEIT Test Predict Outcomes in Organizations?

Does the MSCEIT test predict outcomes for organizations?

A. Sometimes... EI Helps Those Who Must Maintain Relationships in Organizations

☯	Higher-EI employed undergraduates were rated better-performing by supervisors	$r = .22$
☯	Higher customer satisfaction with claims adjusters	$r = .46$, $p < .05$
☯	Preference for a team-oriented management style	$F = 2.96$, $p < .06$
☯	Higher-quality vision statements by groups higher in EI, after controlling for the Big Five personality traits	$p < .05$
Sources: Janovics & Christianson, 2000; Rice, 1999; Moss, 2001; Coté, Lopes, & Salovey, in preparation		

Well, I can't do this quickly, and I'm not sure I have enough time enough to do anything but give you a general idea. When you are lower in the organization and you have to deal with things like supervisors and customers and stuff like that, it really seems to matter. A group of 150 or so undergraduates, who were employed outside the university… Janovics and Christiansen wrote to their supervisors and asked them for supervisee evaluations and gave the students the MSCEIT and there was about a .22 correlation between the undergraduates and their external blind performance rating. In studies of claims adjusters, teams of claim adjusters with higher-than-average EI, this is a small sample study, you can take it with a grain of salt. But they had better rated customer satisfaction, with a correlation of .46.

And another study of, I believe, Yale MBAs, I am not sure about this, in a study of Yale MBA's who were doing sort of simulated or maybe actual consultancy projects, those groups who had higher scores on the MSCEIT, particularly branches 3 and 4, what we call strategic EI, were able to write more motivating and higher quality vision statements than others.

*Q: Does EI Predict Relationships
with Others?*

Does higher EI predict better relationships with others? Does lower EI predict poorer relationships with others?

A: Yes, the MSCEIT Predicted these Positive Aspects of Relationships:

☻	12 month old infants rated as securely attached in an experimental setting, have mothers who score higher on the MSCEIT	$p < .05$
☻	Higher EI scores in adults correlate with higher scores on attachment scales	$r = .28$, $p < .05$
☻	Higher EI scores predict more "objects of attachment" (e.g., photos, letters) around the home	$r = .30$'s, $p < .001$
Sources: Marsland & Likavec, 2003; Kafetsios, 2001; Brackett & Mayer, 2003		

Yes, the MSCEIT predicts a number of positive criteria concerning relationships. Most dramatic to me, that 12 month old infants rated as securely attached in the experimental strange situation have mothers who score higher on the MSCEIT. Higher EI scores in adults correlate with more objects of attachment, like photos and love letters from others and friends around the house.

A: Yes…Positive Aspects of Relationships (Continued):

☯	Higher MSCEIT scores correlate with higher rated social support networks	$r = .15$ to 28, $p < .05$
☯	Higher scores on Strategic MSCEIT branches (understanding and management) predicted: ▪higher quality interactions with friends among 4,500 social interactions in diaries of 100 German college students ▪higher rated social sensitivity in sociometric data among Yale University students	$p < .05$
Sources: Fullam, in preparation; Coté, Lopes, Salovey, & Beers, in press.		

People scoring higher on the MSCEIT are also better liked and are better evaluated by friends and colleagues, as highlighted in this slide.

A. Yes, Low Scores on the MSCEIT Predicted these Negative Aspects of Relationships:

☯ More fights, drug use	$r = .21$ to $.40$, $p < .05$
☯ More alcohol and tobacco use	$r = .15$ to 24, $p < .05$
☯ Higher ratings of aggression by peers at school	$r = .20$ to $.46$, $p < .001$

Sources: Formica, 1999; Trinidad & Johnson, 2001; Rubin, 2000; (N = 48).

And lower EI, this is the first finding that seems to create a pattern with the MSCEIT so I'm very partial to it, lower EI, also predicts the number of negative outcomes more fights and drug abuse, more alcohol and tobacco use and higher ratings of aggression, by peers at school.

Q. What Does EI Look Like in An Individual Case?

As one really quick illustration of how this impacts the individual, consider the following case study. We asked some adolescents to tell us about a problematic social experience that they had had, where their peers wanted them to do something that they felt was conflictual for them.

Highest Scoring Adolescent

- *Participant 6 (16 year-old female; VIQ: 133; EIQ: 128)*: Once my friends wanted to sneak in someone's room and paint them while he slept. It began as joking around ("wouldn't this be funny; could you believe it if?). Then it slowly evolved into dares ("I bet you wouldn't," or "I dare you to.").

This is just one story that a high EI student told. "Once my friends wanted to sneak in someone's room and paint them while he slept, it began joking around," and you'll notice I think it pulls together what a lot of other people are talking about, she is very socially acute, she notices the beginning of the event, the beginning of the emotion. "It began as joking around, 'wouldn't this be funny,' 'could you believe it if,' then it slowly evolved into dare, 'I bet you wouldn't or I dare you to,'" so notice she's not waiting until … she's watching the evolution of the conflict before it gets too big.

Highest Scoring Adolescent (cont.)

I felt like it was betraying the trust I had with the other person, I didn't feel right with sneaking up on a sleeping person with no way to defend himself, and I thought doing this would make the person have his feelings hurt. I know how little pranks like this could really hurt someone's feelings…

Then she says, "I felt it was betraying the trust I had with the other person, I didn't feel right, sneaking up on a sleeping person with no way to defend himself and I thought doing this would make the person have his feelings hurt. I know how little pranks like this could hurt someone's feelings." So here is the identifying emotion but also understanding how an act will bring a certain emotion about.

So we asked how did you handle it, a standard question. "I told my friends straight out, it was a degrading thing to do and they shouldn't be so cruel." And now here in this last portion of the response is that abstract reasoning. We asked how would your parents have reacted, and she said, first "they would have been proud," but then she switches perspective and she says, "they might also have said that I ruined a perfectly harmless joke." And that to me is intelligent, to me she is perceiving it, she is understanding what's going to happen, and she's able to shift perspectives on it and to see it from different vantage points.

General Conclusions about Emotional Intelligence

- EI can be defined as an intelligence
  - to perceive emotion, facilitate thought, understand emotion, and manage emotion
- The MSCEIT serves as a performance measure for EI.
- Results from the MSCEIT suggest that it does measure the capacity to reason with emotion, and of emotion to assist thinking.
- The MSCEIT measures something new relative to the existing and most widely used measures of personality and intelligence.
- High EI relates to stronger attachments; low EI to deviant, and often, problematic, behavior such as drug use and violence, among other findings.

Now to the conclusion; first of all our research owes a debt of gratitude to many of those in this room, the rate of learning about emotional intelligence has been vary fast. The definition, our ability model has evolved and I recommend the 97 definition for you.

As, Alicia Grandey, Kathryn Lively, and Aaron Ben-Ze'ev have pointed out, emotion is important to society, I left that out, I often started with these ideas, but those three contributors made such a good case for it, I didn't need to. As the work of Rosalind Picard and W Lewis Johnson point out, the rules of emotional and social hypothesis can be worked out so carefully that computers can follow them. Those are substantiations that exhibit that we know a great deal about how emotions operate. I'm not saying we know everything, obviously we don't, but we know enough to really make some good inroads here. And as Klaus Scherer points out, appraisal is important to emotional intelligence and Paul Ekman and Maureen O'Sullivan also pointed out, facial recognition is critical to emotional intelligence. Moreover, as Gerard Saucier, William Revelle, and Fabio Sala point out, there are many other personality dimensions of relevance to the individual.

<div style="background:blue;color:white;">

<h1 style="color:yellow;text-align:center;">General Conclusions
about Emotional Intelligence</h1>

- EI can be defined as an intelligence
    - to perceive emotion, facilitate thought, understand emotion, and manage emotion
- The MSCEIT serves as a performance measure for EI.
- Results from the MSCEIT suggest that it does measure the capacity to reason with emotion, and of emotion to assist thinking.
- The MSCEIT measures something new relative to the existing and most widely used measures of personality and intelligence.
- High EI relates to stronger attachments; low EI to deviant, and often, problematic, behavior such as drug use and violence, among other findings.

</div>

I think of emotional intelligence and intelligence as dimensions of personality also, there are other dimensions of personality that are important as well. When we say emotional intelligence is what we say, we also mean it's not optimism and it's not niceness and it's not persistence and it's not zeal and it's not all those other things that can go into character. We're not saying those are bad things, there great things to study, but this is different; a different entity. Also Klaus Shearer, Karl Heider, and Gerald Saucier point out we have to watch for cultural differences and I didn't get into that during this presentation; but we acknowledge that they exist.

As Maureen O'Sullivan points out and as our data confirms, there is no relationship between self-reported EI and actual EI, to measure actual EI you have to use a performance measure. More generally speaking, EI can be defined as an intelligence: the ability to perceive emotions, facilitate thoughts, understand emotion and manage them. The MSCEIT serves as a reasonably good performance measure of emotional intelligence. And it's convenient to use, so we can learn about emotional intelligence and emotionally intelligent people. The MSCEIT measures something new relative to the existing and most widely used measures of personality and intelligence. And finally high EI relates to stronger attachment, lower EI relates to deviant and also problematic behavior such as drug use and violence among other qualities. Thank you very much.

DISCUSSION

DISCUSSION

DR. ROBERTS: We have a bit of time for questions, who would like to ask a question?

DR. GRANDEY: Thank you Jack, that was great, I have a couple of quick questions, one was about you said that there was no relationship between self reported EI and the ability to do EI, correct? That just seems odd to me, because wouldn't people who have a high ability as demonstrated performance measures wouldn't they also have high understanding and awareness of their ability and rate themselves higher, and clearly that doesn't work on the other end, but wouldn't you see a weak relationship because of?

DR. MAYER: I'm sure some would, and I appeal to your experience as a professor to answer this question, and I draw on the analogy of general intelligence where the relationship ranges from 0 to .35 depending on what study you're talking about. You have students who are really bright, and know they are, you have students who are really bright but are so modest or have some damaged self image problem or something who think they are not very bright even if you know they are, and you have students who are a lot lower and think they are intelligent, and you have students that are not so bright and they don't even understand the question so they can go anywhere with it. And that really limits your correlation.

DISCUSSION

DR. GRANDEY: Okay, and I was surprised how few studies that you were able to review about work place outcome and obviously from my background that's what I'm interested in, is that something that you and your group are pursuing with the MSCEIT I mean, and I think partially because it's a very involving and long measure so maybe it's difficult to administer to a working sample?

DR. MAYER: I was going to say something different about it. The test is two years old, it was a pre-release. So now I think we are going to see people bringing that in to the workplace. It takes about, it's a commitment, it takes about 40 minutes or so to administer.

DR. LEN WHITE: I have two questions, one do you have a website we could go to learn more about MSCEIT?

DR. MAYER: I don't myself, my colleague David Caruso has a website which is emotionaliq.org for academic types there is a .com there too, he is more of a commercial type person so he's got two of them.

DR. LEN WHITE: The second question, thank you for permitting two…

DISCUSSION

DR. MAYER: I don't know if you're coming from an academic or from a more a consulting perspective, if you're academic and you're interested in papers, just email me and I will be happy to send you what you like and I will put that email up in a minute, but let me listen to your question.

DR. LEN WHITE: I have no idea which perspective I am [laughter]… the other piece I would like to inquire about is in the talk earlier this morning, we talked, mentioned about training and staying in the area perceiving emotion… what is your thoughts about the trainability about what's being measured here by the MSCEIT?

DR. MAYER: Well, we said that I think from the start that there is no question that you can teach emotional knowledge. And I think Paul's CD is terrific for that, in fact I'm going to run right out and get them. So I can teach myself. We do hope the test is not going to be that easily trained, but we don't know. We like to think by perfecting the content of the test, by not giving away the answers and so forth, that although people will raise their knowledge, which we want them to do, we want them to raise their emotion knowledge, if that helps them (for some people it's not necessary, I think). We think that there's room for an emotional assessment test, an emotional achievement test.

DISCUSSION

DR. KATHERINE WHITE: You seem to have a lot of data on your measure, I'm wondering if you've looked at any, if you've tested your measure with older adults, if not, or if you have and have that data, I'm curious to see if you have the data or your prediction to whether emotional intelligence as an ability might behave more similar to maybe a fluid or crystallized ability across the life span do you have any predictions about that?

DR. MAYER: I don't, because I think about it and I think it could go both ways, I tend to think that a cohort effect that younger people are more interested in emotions than older people are and for that reason they maybe a little bit better at performing in very late years, I think of that more of the cohort effect than necessarily an age effect, I think it's an empirical question I hope people like Derrick Isaacowitz and others who are here, who do research with an older population, would answer that question for us sometime soon.

MS. MACCANN: I just wanted to say I know there is an Australian researcher, who has used the MSCEIT with an older population, with people in their 60's and 70's and they do tend to be of higher emotional intelligence.

DR. MAYER: And that was with the MSCEIT?

MS. MACCANN: That was with the MSCEIT.

DR. MAYER: Is that right!

DR. ROBERTS: OK, that's it for this presentation. Thanks, Jack.

Measuring Knowledge Based Constructs

Pete Legree
& Joe Psotka

MEASURING KNOWLEDGE BASED CONSTRUCTS

DR. ROBERTS: It gives me great pleasure to introduce Peter Legree from ARI, who will discuss "Measuring Knowledge Based Constructs".

DR. LEGREE: My talk will describe conceptualizations that focus on consensual models and their role in knowledge assessment and emergent domains; one of which I think is emotional intelligence, although I think there are probably many more. I should point out that I am heavily indebted to Joe Psotka and Douglas Detterman, with whom I worked for the last 10 years to develop some of these concepts.

Measuring Knowledge Based Constructs

Pete Legree
& Joe Psotka

These concepts relate well to the growing use of situational judgment and scenario based scales tests to assess knowledge and to simulate situations requiring expertise in performance domains such as leadership and sales, as well as to assess emotional and social intelligence, and for some rather specific purposes I think cognitive aptitude. Scales developed for these various purposes are similar in that they present or describe a situation and require examinees to indicate how they might respond or interpret the situation. These responses or interpretations are then analyzed to quantify performance on the tests. While most of these applications utilized expert groups to develop scoring standards (cf. Hedlund et al., 2003), sometimes scoring keys have been constructed by analyzing data collected from large groups of respondents who were knowledgeable concerning the subject domain but could not be qualified as experts (Legree 1995; Legree, Martin & Psotka, 2000; Legree, Heffner, Psotka & Martin, 2003). These individuals might be described as journeymen, but certainly not as certified experts in any of the traditional senses.

Tolstoy

Happy families are all alike; Every unhappy family is unhappy in its own way

Many diverse opinions can be united into a better consensual whole ..many novices can combine to outperform an expert

It takes a village to grow a child

In these applications, the use of non-expert groups to develop scoring standards was termed "Consensus Based Measurement", and the utility of this method is the focus of this discussion. Scoring these scales has relied on the assumption that the scoring keys developed from groups of non-experts closely approximate those that would be from experts, and this concept is implied by this quote from Tolstoy and the African village proverb.

The perspective underlying the quote and the proverb is that knowledge may be distributed across a population of individuals and that while it's possible for individual opinions to be incorrect in numerous ways, opinions can generally be correct in only a single way.

If this idea is generally valid, then consensus based measurement has the promise to expand the spectrum of knowledge that can be addressed in psychological research to include areas for which neither certified experts can be identified nor objective knowledge located; one of which may be emotional intelligence because EI is an example of an emerging domain that largely lacks certified experts and objective knowledge. Furthermore, consensus based measurement has been used to score the Mayer Emotional Intelligence Scale (MEIS; Mayer Caruso & Salovey, 1999) and the Mayer Salovey Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, Caruso & Sitarenios, 2003), which Jack just ran through.

However, the notion that non-experts can be used to develop the "expert" knowledge required to score these instruments may be unappealing to test developers who are not familiar with the strengths and limitations of the approach, and reviewers have questioned its use and assumptions (e.g., Roberts, Zeidner, & Matthews 2001; Zeidner, Matthews, & Roberts, 2001; Schaie, 2001). So it is important to describe consensus-based measurement, and summarize relevant theory and data and discuss limitations of the approach.

Most knowledge tests are based on objective knowledge or on the results of a job or task analysis that associates knowledge with performance in some domain. This approach has proven its worth in many pragmatic areas of personnel selection assessment and development (cf. Anastasi, 1988). Implicit in this approach are expectations that formal and tacit knowledge underlies much performance and observed behavior supports inferences concerning those knowledge attainments.

Construction of knowledge scales traditionally has drawn on either an available and formal knowledge source (such as books written by experts; or pedagogical materials developed over decades of instruction and analysis) or an available pool of experts.

Memory Structures and Knowledge Acquisition : Tulving

Personal Experience

Received Knowledge

Procedural Memory

Episodic Memory

Semantic Memory

Behavior & Performance
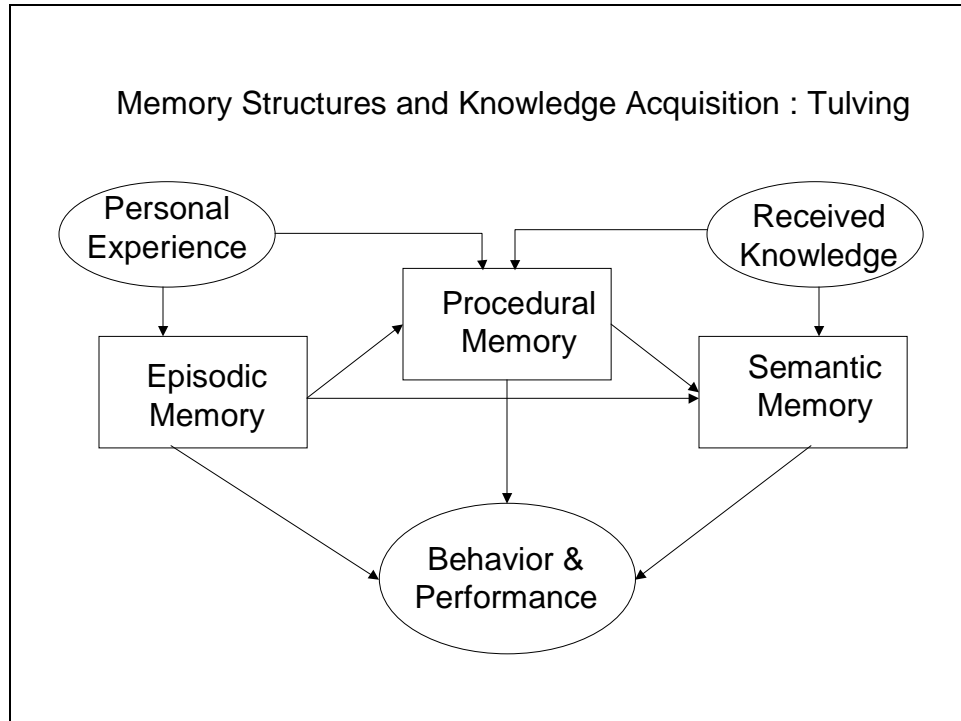
One basic limitation with this approach is that much knowledge is intuitive and tacit, and might be called mere opinion, so there may be no formal knowledge sources, or even experts who can provide appropriate standards. And in some areas, such as art, music, politics, and economics, experts may have or seem to have markedly different views, rationales, and evidentiar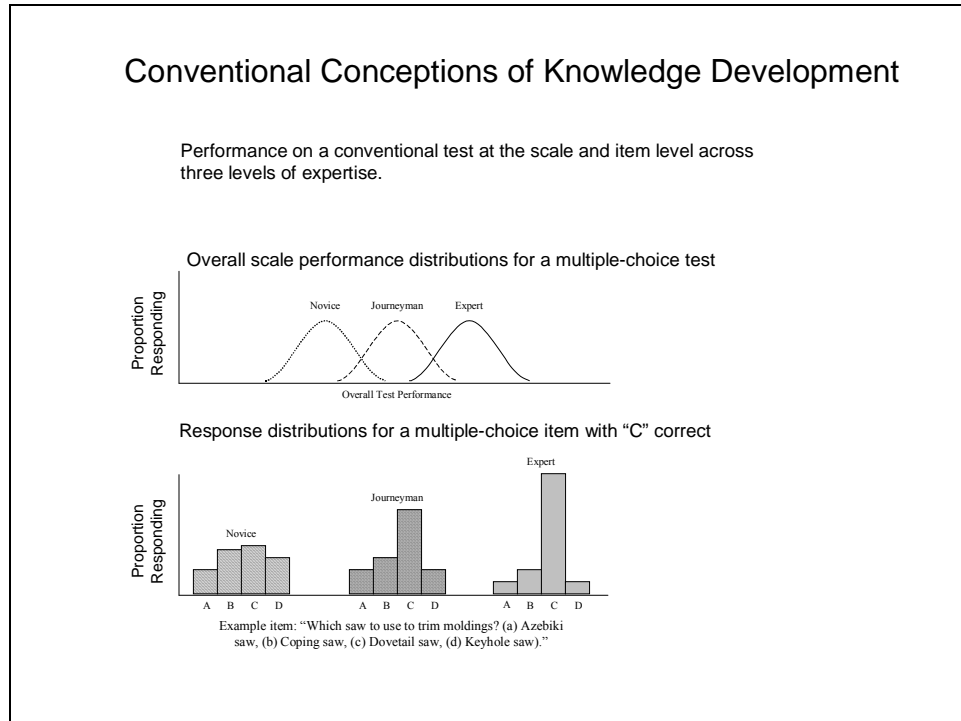y sources than the stratified populations of interest to researchers. Consensus based measurement offers unique analytic powers in these situations, and, as implied, this approach dovetails nicely with the use of Situational Judgment Tests, or scenario, based scales to assess individual abilities.

The belief that much knowledge is experientially based and linked to opinion is rooted in the writings of philosophers (Plato and John Stuart Mill). And the concept can be identified in models of memory, such as Endel Tulving's, that link different forms of memory to different evidentiary sources. Tulving's model is usually used to conceptualize the development of knowledge within individuals, but the model also suggests that different pieces of knowledge may be evidenced across the memory structures of individuals. Thus one aspect of Tulving's model that is often not appreciated is the notion that knowledge might be deduced from understanding the memories and opinions of individuals.

The assessment of knowledge corresponding to soft, emerging domains such as emotional and social intelligence, where the codification and formalization of knowledge is only beginning, cries out for the use of these new technologies. These soft domains are often of considerable consequence: knowledge and expertise related to driving safety, leadership and social functioning can and does substantially impact one's quality of life.

Conventional Conceptions of Knowledge Development

Performance on a conventional test at the scale and item level across three levels of expertise.

Overall scale performance distributions for a multiple-choice test

Response distributions for a multiple-choice item with "C" correct

Example item: "Which saw to use to trim moldings? (a) Azebiki saw, (b) Coping saw, (c) Dovetail saw, (d) Keyhole saw)."
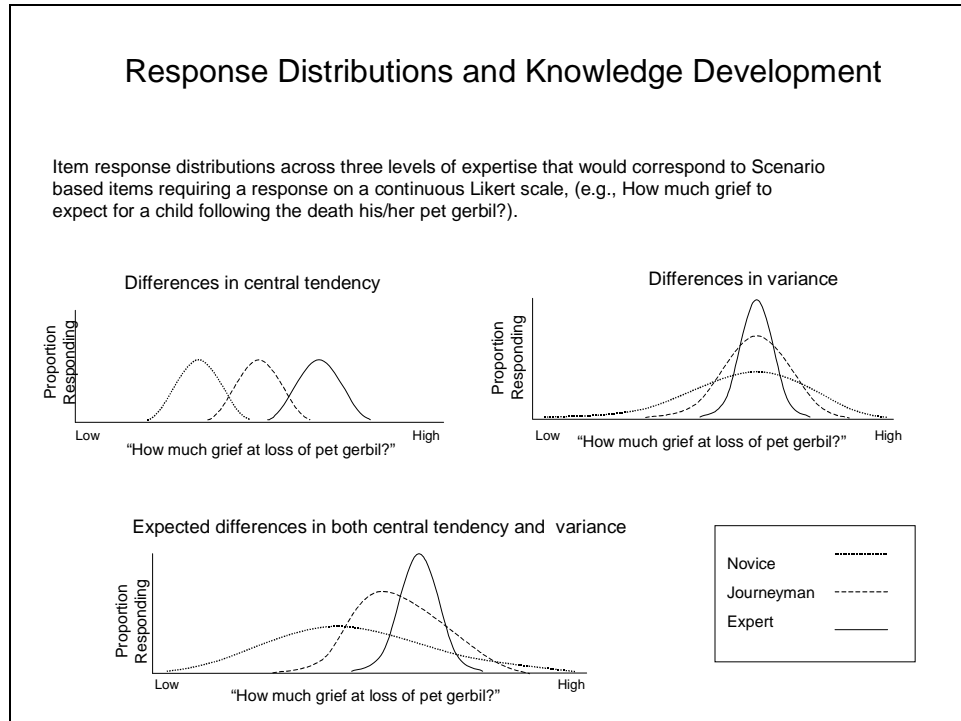
Our conceptualizations regarding consensus based measurement evolved from considering the manner with which item response distributions might change as a function of the expertise of respondent samples. Knowledge is customarily viewed as growing over levels of expertise. If a sample of apprentices was tracked over time, and repeatedly surveyed with standard knowledge items as novices, journeymen and experts, the response distributions described in this figure might be obtained for the test. The distributions in this figure illustrate individual differences as well as the expected increase in knowledge over levels of expertise. For an individual test item, greater proportions of respondents would choose the correct response over increasing levels of expertise, as is illustrated in lower figure (Figure 1b).

This tendency is central to the using examinee data to approximate expert scoring standards. On a more personal level, instructors I've spoken to readily accept the idea that a majority rule algorithm could identify the answers for most multiple choice items. So it seems a little surprising that the approach is not more readily accepted and used. Of course, writing multiple-choice questions generally requires that the correct answer be known, so although this idea is readily accepted, at least by course instructors, it's rarely leveraged.

Response Distributions and Knowledge Development

Item response distributions across three levels of expertise that would correspond to Scenario based items requiring a response on a continuous Likert scale, (e.g., How much grief to expect for a child following the death his/her pet gerbil?).
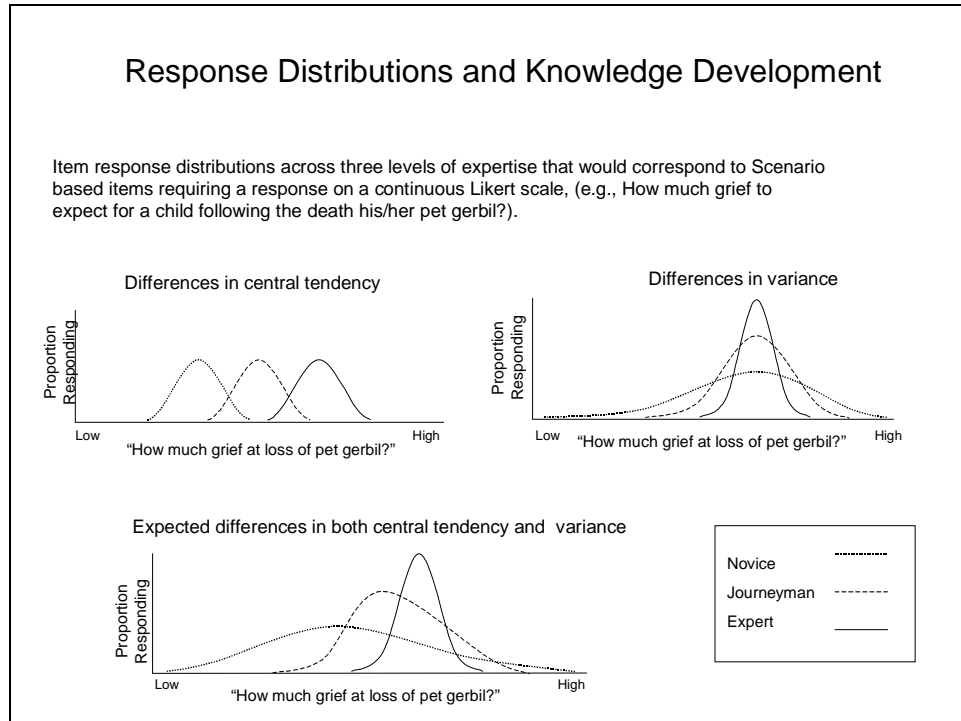
However, suppose examinees were surveyed with scenario-based items that required a Likert rating to be associated with each item. Jack just provided us with five examples; I'll give another. Examinees might be requested to estimate the level or intensity of an emotion associated with some event, such as the level of grief that a child might experience following the death of a pet gerbil. It's important to appreciate that an item with this format contains the correct answer, and can be phrased without knowing the answer. In this example a sample of young children might be thought to approximate a sample of novices. For this type of question, the item response distributions associated with increased levels of expertise, for example young versus older children, might vary in central tendency, or in variance, or in both variance and central tendency.

A change in central tendency might occur as children learn to expect a peer's emotionality to be influenced by discrete events (e.g., when children learn that grief may follow the loss of a family members); changes in the central tendency of these types of response distributions are illustrated in Figure 2a.

A change in variance might occur as children become more accurate in predicting the emotionality linked to these discrete events, in the example, recognizing that a gerbil is just a rodent and its death would be associated with only a moderate level of grief. Figure 1b illustrates a change in variance of response distributions associated with increased accuracy (estimating the level of grief).

Response Distributions and Knowledge Development

Item response distributions across three levels of expertise that would correspond to Scenario based items requiring a response on a continuous Likert scale, (e.g., How much grief to expect for a child following the death his/her pet gerbil?).

Differences in central tendency

Differences in variance

"How much grief at loss of pet gerbil?"

"How much grief at loss of pet gerbil?"

Expected differences in both central tendency and variance

"How much grief at loss of pet gerbil?"

Novice
Journeyman
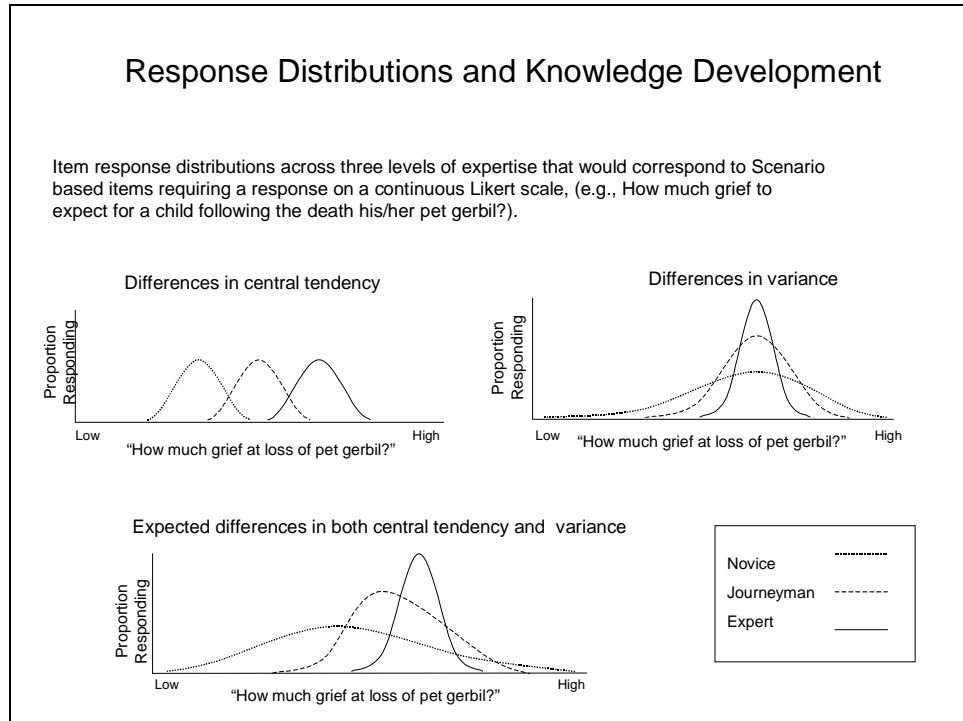Expert

Both these trends have relevance to understanding the growth and refinement of knowledge. Naïve individuals lack any basis for understanding relationships or events, and their responses, often representing poorly formed cognitions, may not be sensible, sometimes indicating ignorance of even basic relationships, and sometimes overstating their importance. But with increasing degrees of sophistication, individuals will become increasingly aware and accurate in their understanding of relationships and events.

This conceptualization suggests that examinees might tend to agree with each other to a greater extent across increasing levels of expertise for both conventional and scenario-based test items. Given this conceptualization, we recognized that estimates of the central tendency of expert response distributions for individual, scenario-based items would be approximated by estimates of the central tendency of non-expert (e.g. journeymen) response distributions when the growth of knowledge over expertise is associated primarily with changes in variance (Figure 2b). This observation also applies to conventional multiple-choice items (Figure 1b), but it is of little practical value because writing multiple-choice items requires that the correct response be known. Scenario based items do not always require that the correct response be specified or even known. In the gerbil example, the test developer need only specify the emotion, (i.e., grief) and allow examinees to respond to the item.

Response Distributions and Knowledge Development

Item response distributions across three levels of expertise that would correspond to Scenario based items requiring a response on a continuous Likert scale, (e.g., How much grief to expect for a child following the death his/her pet gerbil?).

Differences in central tendency

Differences in variance

Expected differences in both central tendency and variance

Novice
Journeyman
Expert

Initially, at least, we expected that changes in response distributions over level of expertise would be associated with changes in both central tendency as well as variance. This model is intermediate and is represented in Figure 1c. This model suggests a recursive procedure could be developed to (A) sequentially identify groups of individuals with increasing levels of performance and (B) use this information to develop valid scoring keys for scenario based test items. These standards would then be referred to as "consensus based standards" and the process as "consensus based measurement".

SJT Conceptualization

- Implicitly or explicitly describe a scenario to simulate or depict an event. The scenarios may be oriented towards responding to problems, maintaining success or interpreting an event based on personal experience and world knowledge.
- Provide a list of alternatives associated with each scenario. The lists may include action-oriented or interpretative alternatives, or may allow the examinee to respond in an open-ended manner to describe his or her opinion and knowledge.
- Obligate examinees to evaluate either the alternatives associated with the scenarios (e.g., rating the appropriateness of the alternatives or identifying the most relevant alternative) or to generate new alternatives and analyses of the scenarios in the case of an open-ended response.

Situational judgment tests (SJT) are relevant for studying changes in item response distributions over expertise, in both central tendency and variance as described above. We describe situational judgment tests broadly as scenario-based scales with characteristics you see on that slide.

This description is intended to be more inclusive then might be discerned by reviewing scales that are often described as SJTs; these scales are typically lists of problem-action choices. How these evaluations are scored against some standard and how these standards are developed is the topic of concern in this essay.

The scoring standards for most existent situational judgment tests are developed by surveying subject matter experts to identify appropriate responses (cf. McDaniel, Morgeson, Finnegan, Campion & Braverman, 2001). A typical procedure involves collecting ratings data for alternatives and using the mean expert ratings to develop a scoring standard. A percent correct agreement, a deviation measure, such as distance score, or a correlation of an examinee ratings with the scoring standard is then computed to evaluate examinee responses in comparison to the expert-based standards. Existing SJTs are potentially ideal for testing the hypotheses I've been suggesting to the extent that both expert and examinee ratings might be obtained and compared for scoring purposes.

SJT example item

1	2	3	4	5	6	7	8	9	10	11

Extremely
Inappropriate

Neither Inappropriate
Nor Appropriate

Extremely
Appropriate

You are a Work Center NCOIC. Over the past several months you have noticed that one of the other Work Center NCOICs in your Flight hasn't been conducting his Common Task Training (CTT) correctly. Although it hasn't seemed to affect the Flight yet, it looks like the Flight's marks for CTT will go down if he continues to conduct the CTT incorrectly. How appropriate are the following actions?

_2_ a. Do nothing since performance hasn't yet been affected.

_7_ b. Have the Work Center NCOIC meeting and tell the Work Center NCOIC who has been conducting training improperly that you have noticed some problems with the way he is training his troops.

_8_ c. Tell your Flight sergeant about the problem.

_10_ d. Privately pull the Work Center NCOIC aside, inform him of the problem, and offer to work with him if he doesn't know the proper CTT procedure.

Here is an SJT item developed to evaluate supervisory knowledge within the Army. This test item was usually scored as using a variation of a percent correct procedure in which the examinee identifies the most and least appropriate action. But the item could also be scored as a distance by comparing the responses of examinees and a scoring standard as shown.

Consensus Based Measurement Expectations

1. High correlations between scoring standards
based on expert and examinee data
2. High correlations between scores based on
these standards
3. Adequate criterion relevance

So given our conceptualizations regarding the changes in response distributions over levels of expertise, and these characteristics of scenario-based scales, these following relationships might be expected for scales scored with a census based approach. Assessing the correspondence between expert and examinee scoring standards requires that similar data be collected and that the expert sample be well described. We know of seven databases that have implications for these expectations, although there are probably many more. It's important to recognize that most of these data do not have direct relevance to emotional intelligence.

The value of this information is that the diversity of applications addresses the general utility of the procedure. I think the justification for these expectations is so strong that their assessment is better described as a parameter estimation exercise. So let's review these data for implications for these conceptualizations.

Supervisor and Social Intelligence SJT Data

- Three SJTs were administered to 200 AF recruits to assess supervisor knowledge, social interactions and social insight.
- For the Supervisory Scale, Score Standard convergence: r=.74 (.95)
- Score Convergence Using in AF and Army standards: r=.88
- Factor structure supported a g-loaded model of social intelligence

R's	Quartile 2	Quartile 3	Quartile 4	AF Overall	Army	Mean Item SD
Quartile 1	.96	.92	.71	.98	.76	2.1706
Quartile 2		.93	.76	.98	.73	2.3021
Quartile 3			.76	.98	.73	2.4197
Quartile 4				.80	.49	2.5281
AF Overall					.74	2.3759

We first evaluated this conceptualization of knowledge development by comparing expert based scoring standards that reflected the opinions of a small number of subject matter experts (i.e., mean expert ratings) and the mean ratings for the items as computed across examinees. That Supervisor SJT described 49 scenarios and listed a total of 202 alternatives. Each scenario described an interpersonal problem and presented alternatives as possible solutions to the problem. The test had been developed using Senior NCO's as experts who rated each of the alternatives and was intended for administration to first term Army soldiers using a Most-Least format. However, the scale was also administered to a sample of AF recruits using the ratings format. Thus it was possible to describe the effectiveness ratings distributions for sample of NCOs and fairly novice recruits.

The correlation between the expert-based scoring standard and the examinee mean item ratings was .74, (N=198, p<.001), and a very high parameter, above .95, was estimated by correcting the observed correlation for attenuation of the reliability of each set of observations (i.e., the mean expert and examinee ratings). The scores defined by the two standards correlated .88. In scoring these scales response distributions were equated by standardizing within subject. So these initial data, at least, support the conceptualization of expertise being associated with decreased variance.

Supervisor and Social Intelligence SJT Data

- Three SJTs were administered to 200 AF recruits to assess supervisor knowledge, social interactions and social insight.

- For the Supervisory Scale, Score Standard convergence: r=.74 (.95)

- Score Convergence Using in AF and Army standards: r=.88

- Factor structure supported a g-loaded model of social intelligence

R's	Quartile 2	Quartile 3	Quartile 4	AF Overall	Army	Mean Item SD
Quartile 1	.96	.92	.71	.98	.76	2.1706
Quartile 2		.93	.76	.98	.73	2.3021
Quartile 3			.76	.98	.73	2.4197
Quartile 4				.80	.49	2.5281
AF Overall					.74	2.3759

We tried to revise and improve the examinee-scoring standard by using a program to progressively select more knowledgeable examinees to produce a "smart examinee" scoring key, However only a marginal gain was realized by refining examine-based standards by using only higher scoring examinees to define the standard. This is because standards defined using selected samples of examinees were highly correlated with the inclusive examinee standard. For example standards based on the top quarter and the total sample correlated .98, which admittedly is a part-whole correlation. But it may be more surprising that when the data were divided by quartile, even the standard defined by lowest performing examinees correlated .72 with the top quartile and .49 with the NCO standard. Standards based on the top 3 quartiles correlated between .92 and .96 and correlated between .72 and .76 with the NCO values. These correlations confirmed that the mean ratings of examinees might provide an alternate-scoring standard for the SJT, and this realization raised issues concerning the appropriateness of the two standards. From a reliability perspective, the examinee-based standard may be viewed as preferable because the values based on the opinions of a large sample of examinees (193) can be expected to be reliable, and because the ten years had elapsed between the collection of the NCO and AF data might allow shifts in social standards.

Supervisor and Social Intelligence SJT Data

- Three SJTs were administered to 200 AF recruits to assess supervisor knowledge, social interactions and social insight.

- For the Supervisory Scale, Score Standard convergence: r=.74 (.95)

- Score Convergence Using in AF and Army standards: r=.88

- Factor structure supported a g-loaded model of social intelligence

R's	Quartile 2	Quartile 3	Quartile 4	AF Overall	Army	Mean Item SD
Quartile 1	.96	.92	.71	.98	.76	2.1706
Quartile 2		.93	.76	.98	.73	2.3021
Quartile 3			.76	.98	.73	2.4197
Quartile 4				.80	.49	2.5281
AF Overall					.74	2.3759

A confirmatory factor analysis of scores for this SJT, two additional social intelligence scales, which could only be scored with a consensus based procedure, and a set of conventional cognitive aptitude tests, demonstrated the existence of a separate factor corresponding to a model of social intelligence. However the possibility that the results primarily reflected a method effect could not be discounted with this database. But at least for this database, and as far as the expectations for the scoring standards were concerned, most of the consensus based scoring expectations held.

Unobtrusive Knowledge Tests & Tacit Driving Knowledge Scales

The UKT scales were designed to measure psychometric g, while the Driving Knowledge scales were designed to measure safe driving knowledge. The Driving Knowledge scales were expected to have low g-loadings because meta-analyses indicated a .10 correlation between g and crash involvement . For these scales:

- Expert standards were not available and all scales were scored consensually
- Factor scores extracted from the UKT and the ASVAB correlated .54 (corrected for range restriction .80)
- Factor scores extracted from the Driving Knowledge tests correlated with crash involvement criteria (.19) while having very low loadings on g, .11 (corrected for range restriction .16).
- Only modest improvement was associated with scoring refinements

In additional work, consensus based measurement was assessed by validating scales based on consensual measurement against conceptually important criteria. Specifically, we developed and validated two sets of scales: six Unobtrusive Knowledge Tests (UKTs) to measure general cognitive ability, and two Tacit Driving Knowledge Tests to assess knowledge related to driver safety.

These scales were intended to have practical utility. The UKT battery was developed as a tool to estimate g within anonymous Internet and paper-based surveys and thereby provide market segmentation information. The UKTs were designed to appear to be opinion questions and not knowledge test items and by intent it was not possible to use conventional reference materials to answer the items. For example, one scale required individuals to estimate the frequency of various words in the spoken English. Thus performance on these scales could only be evaluated using consensus based scoring algorithms.

The driver safety scales were intended to provide information that might lessen crash involvement. This latter domain is unusual because meta-analysis indicates only very low relationship between g and crash involvement. Thus this approach was applied to assess knowledge in two domains that differed greatly in their loadings on general cognitive ability.

Unobtrusive Knowledge Tests & Tacit Driving Knowledge Scales

The UKT scales were designed to measure psychometric g, while the Driving Knowledge scales were designed to measure safe driving knowledge. The Driving Knowledge scales were expected to have low g-loadings because meta-analyses indicated a .10 correlation between g and crash involvement . For these scales:

- Expert standards were not available and all scales were scored consensually
- Factor scores extracted from the UKT and the ASVAB correlated .54 (corrected for range restriction .80)
- Factor scores extracted from the Driving Knowledge tests correlated with crash involvement criteria (.19) while having very low loadings on g, .11 (corrected for range restriction .16).
- Only modest improvement was associated with scoring refinements

These measures required individuals to respond to items using Likert scales, for example, estimating the frequency of words and terms used in oral communication or the extent to which drivers should moderate speed when confronted with driving hazards. Construction of these scales leveraged conceptualizations of incidental learning and tacit knowledge to predict and understand human performance. This type of knowledge and associated expertise is usually acquired slowly and incrementally as a result of experience and reflection upon those experiences in a manner consistent with Tulving's memory model and as has been discussed by Sternberg (Sternberg et al., 2000). For these scales, neither an objective knowledge base nor experts could be identified to develop scoring standards.

The UKT battery was administered to a sample of Air Force recruits. Factor scores extracted from this experimental battery correlated .54 with factor scores extracted from a conventional test battery (i.e., psychometric *g*); with a .80 estimate based on the range restriction correction (Legree, Martin & Psotka, 2000). Five of the six experimental scales also correlated significantly with psychometric *g*. This parameter estimate, .80, is typical of correlations obtained among cognitive ability test batteries (cf. Carroll, 1993). A Lisrel analysis of the corrected correlation matrix estimated a .97 path coefficient between the two latent factors corresponding to the Unobtrusive Knowledge and Conventional Test Batteries.

Unobtrusive Knowledge Tests & Tacit Driving Knowledge Scales

The UKT scales were designed to measure psychometric g, while the Driving Knowledge scales were designed to measure safe driving knowledge. The Driving Knowledge scales were expected to have low g-loadings because meta-analyses indicated a .10 correlation between g and crash involvement . For these scales:

- Expert standards were not available and all scales were scored consensually
- Factor scores extracted from the UKT and the ASVAB correlated .54 (corrected for range restriction .80)
- Factor scores extracted from the Driving Knowledge tests correlated with crash involvement criteria (.19) while having very low loadings on g, .11 (corrected for range restriction .16).
- Only modest improvement was associated with scoring refinements

The tacit driving knowledge tests were administered to Army soldiers and correlated significantly with crash involvement data, -.11 to -.20 (Legree, Heffner, Martin, Psotka, & Medsker, 2002), and as expected, had only low correlations with measures of cognitive ability. While modest, these values generally exceed coefficients obtained for stable characteristics and they carry implications for improving driver safety.

To summarize, these two applications demonstrate that it is possible to develop and validate scales with widely varying g loadings without using subject matter experts or objective knowledge, instead using consensus based measurement. Thus the data confirm the method as well as validate the scales.

Additional data with expert and examinee based standards and scores

Scale / Source	N: Expert/ Examinee	Scoring Key convergence (r)	Score convergence (r)
NCO21 Supervisory SJT (Porr)	88/1891	.89	.95
MSCEIT (Mayer Caruso & Salovey)	21/2112	90	98
Management SJT (Motowidlo)	25/400	.79	--
TKML (Psotka & Sternberg)	50/355	.96	1.00

The above data demonstrate the efficacy of the approach, especially in terms of producing predictive validity, and by implication in terms of producing useful scoring standards. There is little doubt that consensus based measurement can be used to score tests developed for these unusual soft knowledge domains that lack formal sources of knowledge, which either may be very highly g-loaded, or have very minimal g loadings. It was a pleasant surprise that the use of a recursive-scoring algorithm was not necessary for any of these applications as originally expected, and this reduced the computational complexity of the approach. The remaining four datasets address the generality of the finding of agreement between expert and consensus based scores.

The largest of these data sets corresponds to the Non-commissioned Officer (NCO) SJT developed to evaluate supervisory skills for senior enlisted soldiers. The NCO SJT described 71 problem scenarios and listed 362 actions. To evaluate consensus based scoring, response protocols were scored using both expert (N=88) and consensus (N=1891) based standards (Heffner & Porr 2000; Porr, personal communication, July 2003). Overall performance scores correlated .95 and scoring standards correlated .89.

Additional data with expert and examinee based standards and scores

Scale / Source	N: Expert/ Examinee	Scoring Key convergence (r)	Score convergence (r)
NCO21 Supervisory SJT (Porr)	88/1891	.89	.95
MSCEIT (Mayer Caruso & Salovey)	21/2112	90	98
Management SJT (Motowidlo)	25/400	.79	--
TKML (Psotka & Sternberg)	50/355	.96	1.00

Correspondence between expert and consensus based scores has also been demonstrated for the MSCEIT (Mayer at al., 2003), which is arguably the best-developed performance emotional intelligence batteries. The expert group corresponded to 21 members of the International Society for Research on Emotions, and the consensus scores corresponded to 2112 examinees who completed the scale. The correlation between the scores based on the two sets of standards was .98 and the score standards correlated .91.

Additional data with expert and examinee based standards and scores

Scale / Source	N: Expert/ Examinee	Scoring Key convergence (r)	Score convergence (r)
NCO21 Supervisory SJT (Porr)	88/1891	.89	.95
MSCEIT (Mayer Caruso & Salovey)	21/2112	90	98
Management SJT (Motowidlo)	25/400	.79	--
TKML (Psotka & Sternberg)	50/355	.96	1.00

The last database corresponds to the Tacit Knowledge for Military Leadership (TKML) scale (Hedlund et al., 2003; Psotka et al., In Press). The TKML was designed to measure the practical, action-oriented knowledge that Army leaders typically acquire from experience. The TKML was developed with the idea that an ordered hierarchy of expertise in Military Leadership can be created by using the scores of Colonels as a standard and comparing them with U.S. Military Academy (West Point) Cadets, and U.S. Army Lieutenants, Captains and Majors. The scale was administered to groups of soldiers including: 355 Cadets, 125 Lieutenants, 117 Captains, 98 Majors and 50 Colonels. The Colonels comprised the expert group, and this group contains the highest ranking soldiers and those who have served longest in the military (with an average of 18 years service). Comparisons of the consensus based cadet (355 cadets) and expert (50 colonels) scoring standards and scores provide very consistent results with the earlier data. The two sets of score standards correlated .96, and the cadet scores computed using those two standards correlated 1.00. Similar results were found by analyzing the data for the intermediate (lieutenant, captain and major) groups. Despite these correlations, rank was monotonically related to performance on the TKML such that higher ranking soldiers performed on average better than lower ranking soldiers at all levels.

TKML Ratings: SDs by Means
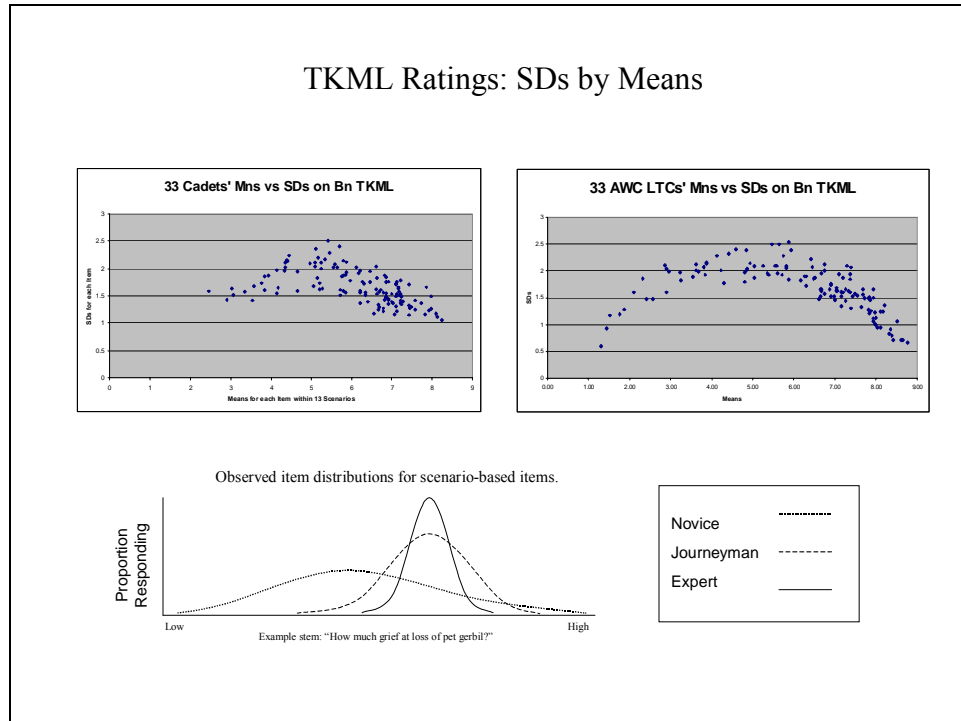
Differences between scoring standards computed using examinee samples can be demonstrated using the TKML dataset, but only by isolating a group associated with a very low level of expertise and comparing their means with values based on the other groups. These relationships are comparable to those summarized for the Project A SJT database. For the top 25% of the Cadets, the correlation with the experts is 0.95. But for the bottom 25%, the correlation with the experts is 0.85. So by artificially restricting the examinee sample to the lowest quartile of the cadet sample, substantial changes in the standards start to emerge, but even then the correlation is still .85.

An interesting aspect that's associated with TKML data is that although the two sets of means are highly correlated, there are differences between the expert and examinee data such that experts agree to a greater extent when identify very poor responses. This is apparent by inspecting the scatter plots for these two distributions.

These findings suggested that modification to the consensus based measurement model might be warranted such that the principle difference between journeymen and experts is represented in terms of increasing accuracy, or from the perspective of item response distributions, decreased variance around the item means. The transition from novice to journeyman may still be associated with shifts in response distributions and means because novices have no or little basis for their responses and their responses would be random or nearly random.

TKML Ratings: SDs by Means

Observed item distributions for scenario-based items.
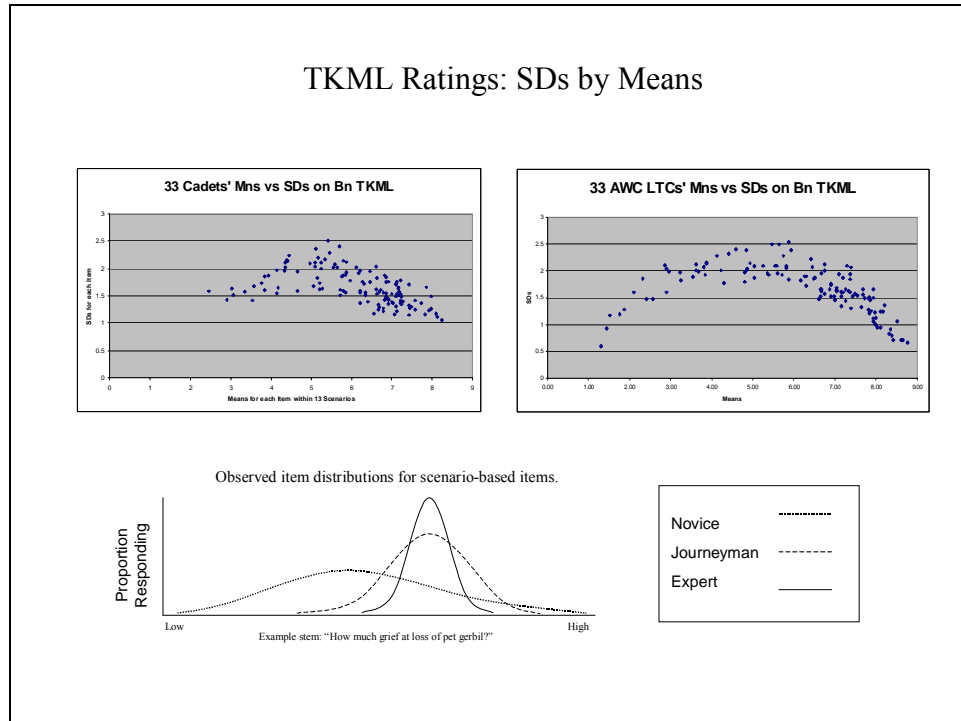
We expect the scatter plot differences reflect censoring in the experiences available to the experts in comparison to the students.

The cadets do relatively well on picking out the good alternatives, but appear to have real difficulty discerning how bad the irrelevant actions are.

They think the bad items are better than the experts do, and their standard deviations (SDs) for the bad items are much higher than the experts'.

As in Tolstoy's magnificent phrase, there are so many different ways to go wrong that they appear to be overwhelmed by the enormity of all the possibilities.

Also, our personal experiences and formal knowledge learned from others are full of advice and experience about the right ways of doing things, but impoverished in the experience of what not to do.

Conceptualizing Consensus Based Measurement : Summary

- Expert and Examinee standards will be highly correlated when the distributions of precipitating events are similar for these groups,
- This is equivalent to the expectation that when exposure to experiences and declarative knowledge is similar over levels of expertise, $r(expert,truth) > r(journeyman,truth) > r(novice,truth)$
- Disagreement is expected when the distributions of exposure to either declarative knowledge or experiences is censored
  - Declarative Knowledge Expectations: Alcohol and crash involvement, Urban crime
  - Proscribed Experiences: Teen Smoking, Sexuality & Alcohol expectations

---

To understand consensus based scoring, it is useful to consider that for most knowledge domains, and especially for procedural knowledge domains, knowledge accumulates as the result of experience (cf. Anderson & Lebiere, 1998). As a greater range of events is experienced, greater levels of knowledge and associated skills will be acquired, and reactions to a new event or situation may reflect increasing levels of sophistication. It follows that the same reasoning applies to exposure to declarative knowledge as outlined by Tulving, although exposure to declarative knowledge in academic settings may be much more systematic while exposure to declarative knowledge through mass communications may be much more repetitive and possibly biased.

In the absence of systematic biases in exposure to experiences or declarative knowledge, components of the novices' thinking should be in error in different ways, but the components that are on the road to expertise should be similar.

When presented with a situation to analyze, novices will have little basis for their opinions, and they will frequently disagree among themselves as well as with experts. Disagreement among novices is expected because the knowledge and cognitive structures associated with an individual novice will reflect either the action of a few unique experiences or the actions of experiences that have marginal relevance to the presented situation. Thus novices will reference different experiences and expectations, and their opinions will tend to be inconsistent, both among themselves as well as with experts.

Conceptualizing Consensus Based Measurement : Summary

- Expert and Examinee standards will be highly correlated when the distributions of precipitating events are similar for these groups,
- This is equivalent to the expectation that when exposure to experiences and declarative knowledge is similar over levels of expertise, r(expert,truth)>r(journeyman,truth)>r(novice,truth)
- Disagreement is expected when the distributions of exposure to either declarative knowledge or experiences is censored
    - Declarative Knowledge Expectations: Alcohol and crash involvement, Urban crime
    - Proscribed Experiences: Teen Smoking, Sexuality & Alcohol expectations

---

In contrast, experts will generally have well-developed, mature knowledge structures reflecting broad, extensive sets of experiences and exposures to declarative knowledge.   While each expert will have a slightly different set of experiences, these sets will largely overlap across individual experts, and with increasing levels of expertise, knowledge structures and related opinions will become progressively more consistent. Journeymen with partially developed and varying levels of expertise will agree at a moderate level both among themselves and with experts, and this moderate level of agreement is based on developing cognitive structures that reflects a modest but not extensive array of experience.

Because procedural knowledge is experientially based and because these experiences are generally dependent on the occurrence of real-world events, various journeymen may have different types of experiences and knowledge, although much of this knowledge will be most relevant to those situations that frequently occur. It follows that the breadth of experience associated with a single expert, while more extensive than that of an individual journeyman, will often be exceeded by the variety of experiences associated with a substantial number of journeymen. The implication of this view for consensus based measurement, as well as for other knowledge engineering applications, is that more information might be present in the knowledge structures of a large number of journeymen than a small number of experts.

Conceptualizing Consensus Based Measurement : Summary

- Expert and Examinee standards will be highly correlated when the distributions of precipitating events are similar for these groups,
- This is equivalent to the expectation that when exposure to experiences and declarative knowledge is similar over levels of expertise, r(expert,truth)>r(journeyman,truth)>r(novice,truth)
- Disagreement is expected when the distributions of exposure to either declarative knowledge or experiences is censored
    – Declarative Knowledge Expectations: Alcohol and crash involvement, Urban crime
    – Proscribed Experiences: Teen Smoking, Sexuality & Alcohol expectations

---

From a mathematical perspective, the correlation of knowledge between individual A and individual B can be conceptualized as the product of the correlation of individual A with the 'truth' and of individual B with the 'truth'. As individuals A and B become more knowledgeable and their opinions more "truthful", their opinions and responses will become more highly correlated (cf. Romney & Weller, 1984).

For many domains progression is reasonable, but for some domains exposure to experiences or knowledge may be heavily censored. This might occur when certain activities are proscribed from segments of the population, for example when youth develop unrealistic expectations concerning alcohol or social interactions, when knowledge that is repeatedly conveyed through mass-communications or within academic settings is one-sided or biased.

These learning theories are most relevant to understanding consensus based measurement when cognitive structures and related knowledge reflect the experience of largely unpredictable events, as does much procedural and tacit knowledge. In contrast, academic knowledge reflects more formal instruction, which is often structured to provide a systematic, highly ordered set of experiences based on objective information, and the surveying of students on topics not yet covered is unlikely to identify much information. However most of these domains correspond to procedural knowledge, be it incidental or tacit, as do many SJTs (cf. McDaniel et al., 2001), and we suspect, many soft, poorly defined domains.

317

Conceptualizing Consensus Based Measurement : Summary

- Expert and Examinee standards will be highly correlated when the distributions of precipitating events are similar for these groups,
- This is equivalent to the expectation that when exposure to experiences and declarative knowledge is similar over levels of expertise, r(expert,truth)>r(journeyman,truth)>r(novice,truth)
- Disagreement is expected when the distributions of exposure to either declarative knowledge or experiences is censored
    - Declarative Knowledge Expectations: Alcohol and crash involvement, Urban crime
    - Proscribed Experiences: Teen Smoking, Sexuality & Alcohol expectations

Thus cognitive theories related to the acquisition of procedural knowledge support the contention that the opinions of a large number of journeymen can be used to approximate those opinions of a smaller number of experts for these types of domains, and this notion is the heart of consensus based measurement.

Conceptualizing Consensus Based
Measurement : Implications

- Supports scale development in emerging domains often lacking
  experts, such as Emotional Intelligence
- Provides economy to test development
- Explicitly invokes the concept of disagreement to understand
  knowledge structures

Much knowledge represents the convergence between many perspectives and truth is commonly believed to exist at the intersection of these perspectives. The perspective that knowledge is rooted in widely diverse opinion is reflected in Tolstoy's observation that "Happy families are all alike; every unhappy family is unhappy in its own way", and from a cross-cultural perspective, the African proverb, "It takes a village to raise a child." The success of these institutions and the relevance of these statements reflect the notion that knowledge can be distributed over individuals, and is consistent with use and development of technologies to identify this type of knowledge and its evidentiary sources for emerging fields such as social and emotional intelligence.

Consensual scoring has several important implications for studying individual differences. First, the approach allows the construction and scoring of scales for knowledge domains for which experts do not exist, or cannot be easily identified. This allows an expansion of the domains for which knowledge tests may be developed, an expansion beyond traditional formal domains into everyday knowledge areas that are meaningful and important in our daily lives. Thus consensus based scoring allows the assessment of knowledge domains which have not been traditionally addressed in psychological or educational research, and broadens the domain of psychological assessment and intelligence research into horizontal aspects of intelligence, one of which may be emotional or social intelligence. This perspective is consistent with theories of implicit and tacit knowledge acquisition and relates well to conceptualizations of social knowledge.

Conceptualizing Consensus Based Measurement : Implications

- Supports scale development in emerging domains often lacking experts, such as Emotional Intelligence
- Provides economy to test development
- Explicitly invokes the concept of disagreement to understand knowledge structures

A second important implication is that consensus based measurement provides economy to test development. The approach allows questions to be posed, answered and scored without the correct responses known a priori. Thus the scale development cycle is shortened because expert responses are not required to construct scoring standards. In addition, costs associated with the production of scoring standards and rubrics are minimized because expert judgments can be expensive to collect while the examinee data are incidental to scale administration. And the use of Likert scales allows distances to be computed at the item level, thus providing ratio as opposed to dichotomous information for each item.

Third, consensus based scoring has the potential to allow the same protocol to be scored against multiple standards. This approach could be useful in studying controversial domains associated with groups that may adopt different perspectives. This approach might relate well to understanding controversial views differing over gender, political affiliation, race, age or sexual orientation or in identifying the basis for competing theories to explain some phenomenon. This application has similarity to multi-dimensional scaling concepts, but the existence of different scoring standards would imply different types of knowledge structures and evidentiary sources correspondent to a shared topic. So the approach invokes the notion of disagreement and inconsistency in the coherence of knowledge structures. Ill-defined domains are characterized by disagreement even among experts.
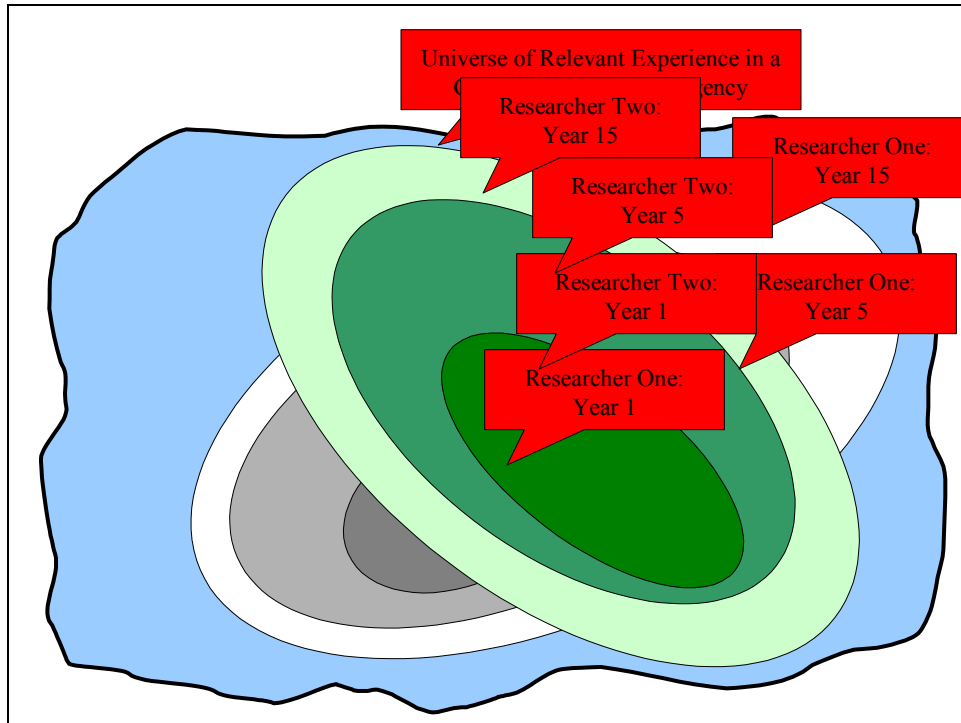
To describe consensus based measurement and to summarize data describing its effectiveness and utility were two goals of this discussion. But the initial model was more descriptive than theoretical, and the concept that expert knowledge can be approximated by surveying large numbers of non-experts must have some limitations. To understand consensus based scoring, it is useful to consider that for most knowledge domains, and especially for procedural knowledge domains, knowledge accumulates as the result of experience (cf. Anderson & Lebiere, 1998). As a greater range of events is experienced, greater levels of knowledge and associated skills will be acquired, and reactions to a new event or situation may reflect increasing levels of sophistication. It follows that the same reasoning applies to exposure to declarative knowledge as outlined by Tulving, although exposure to declarative knowledge in academic settings may be much more systematic while exposure to declarative knowledge through mass communications may be much more repetitive and possibly biased.

In the absence of systematic biases in exposure to experiences or declarative knowledge, components of the novices' thinking should be in error in different ways, but the components that are on the road to expertise should be similar.

When presented with a situation to analyze, novices will have little basis for their opinions, and they will frequently disagree among themselves as well as with experts. Disagreement among novices is expected because the knowledge and cognitive structures associated with an individual novice will reflect either the action of a few unique experiences or the actions of experiences that have marginal relevance to the presented situation. Thus novices will reference different experiences and expectations, and their opinions will tend to be inconsistent, both among themselves as well as with experts.

In contrast, experts will generally have well-developed, mature knowledge structures reflecting broad, extensive sets of experiences and exposures to declarative knowledge.   While each expert will have a slightly different set of experiences, these sets will largely overlap across individual experts, and with increasing levels of expertise, knowledge structures and related opinions will become progressively more consistent. Journeymen with partially developed and varying levels of expertise will agree at a moderate level both among themselves and with experts, and this moderate level of agreement is based on developing cognitive structures that reflects a modest but not extensive array of experience.

Because procedural knowledge is experientially based and because these experiences are generally dependent on the occurrence of real-world events, various journeymen may have different types of experiences and knowledge, although much of this knowledge will be most relevant to those situations that frequently occur. It follows that the breadth of experience associated with a single expert, while more extensive than that of an individual journeyman, will often be exceeded by the variety of experiences associated with a substantial number of journeymen. The implication of this view for consensus based measurement, as well as for other knowledge engineering applications, is that more information might be present in the knowledge structures of a large number of journeymen than a small number of experts.

<div style="border:1px solid black">

DISCUSSION

</div>

DISCUSSION

     DR. ROBERTS: Questions.

     DR. STRICKER: I have been mediating over what seems to me as psychometric paradox underlined consensus scoring, namely when you do the consensus scoring you hope that nearly everybody or certainly a lot of people will agree on a particular response.

     DR. LEGREE: Not really, I think what you're hoping on, what you're hoping for is that your disagreements will be inconsistent.

     DR. STRICKER: Maybe that's the answer cause because I was going to say then, when you then administer that test operationally to the same kind of people who you use for your consensus judgments, you'd expect that most of those people would give the same answer in which case there wouldn't be much variation in the items and wouldn't be much in the way of individual differences, but I think you answered the question.

     DR. LEGREE: I think I do, I can repeat the answer. They're scored as distances from the means, so lower levels of expertise was associated with greater variance around each mean.

<div style="border: 1px solid black; text-align: center; padding: 2em;">

DISCUSSION

</div>

     DR. SAUCIER: Let me ask a practical question, I have another abstract one that I'll deal with if we have time. Jack Mayer was using consensus based scoring and in our exercise he was identifying modal responses then you got a basically a plus score for it if you had the modal response and you didn't get any points I guess if you didn't have the modal response and that was based on some kind of Likert scale? When people use Likert scales some people tend to use the extreme to the scale, some people tend to use the middle, some people like to go high, some people like to go low. So my question is,

     DR. LEGREE: You'd like to know how you'd adjust for that?

     DR. SAUCIER: Yeah, well I'm wondering if it's actually better to score them by looking at the correlation to the consensus judgment rather than identifying how close it is to the modal response.

     DR. LEGREE: That's exactly how the TKLM was scored. Depending upon the application we've transformed within subjects responses so that the responses for all the individuals have the same mean and standard deviation. In general when we've run analysis using simple distances or distances computed using the transform values there is a very high correlation. We tend to obtain slightly higher coefficients in terms of factor analysis and validity data with the more sophisticated approach which really piggy backs on your question, and yes that is correct and there is a solution.

DISCUSSION

DR. KYLLONEN: I have a question: it seems that your procedure will always work unless there is some kind of bias.

DR. LEGREE: I'd like to think so.

DR. KYLLONEN: Unless there is some sort of bias that experts or novices see the world in a different way than experts do.

DR. LEGREE: And I think that is correct.

DR. KYLLONEN: And so I guess the question is this. Is there a diagnostic procedure for you to use so you'd be able to detect whether there might be a systematic bias among novices? Or among the population that would be different from experts without collecting all the data and going through the tedious process of comparing the two scoring keys?

DISCUSSION

DR. LEGREE: Well, I presume there are two variations on your question. One addresses item level statistics, in our analysis we have found occasionally poor items, I will give you an example, in one particular scale we asked, it was actually the passive driving safety scales, passive driving knowledge scales, we asked examinees to estimate the proportion of crashes which involved drunk drivers and the response distribution simply wasn't reasonable. The medium response was 75% and some small portion of response indicated all crashes involved drunk drivers. And it came as a real surprise, a very non sensible response, what we found out was that prior to any long weekend there is a mandatory safety briefing given to all soldiers and almost as a matter of routine, officers hit drunk driving so there is one message that is told at every safety briefing is don't drink and drive, they don't mention possible relationships between being tried and driving or being angry and driving. As a consequence, the examinees overestimated the involvement of drunk drivers and crashes and the item acted very poorly in the factor analysis. In terms of the other questions, suppose there was a domain which was entirely biased, in the sense of comparing the experiences of experts and novices. That is possible, and we haven't worked out methods to identify those domains.

<div style="border:1px solid black; text-align:center;">

DISCUSSION

</div>

DR. LEN WHITE: There seems to me that, and maybe that's what you're getting at; does your data suggest that we may need to be more careful about who we call an expert, or who is called an expert? How does one know if you have an expert?

DR. LEGREE: I think it would be wise to be more careful in terms of who's identified as an expert, but for many domains I'm not sure it really matters. There is a natural progression from apprentice to journeyman to expert and for the most part, experts are simply more accurate and more sensible in their understanding and expectations.

DR. LEN WHITE: Apparently they're people that have been around longer.

DR. LEGREE: That's not always a good thing.

DR. LEN WHITE: It's just interesting.

DR. LEGREE: But it's a good point.

DISCUSSION

PARTICIPANT: I'm referring back in my notes to Karl Heider's definition of culture: "Learned shared ideas about behaviors". So my question is, that I get this contrasted with MH shared ideas about behaviors. Anyway my question is: is the consensus actually culture--is that what you're capturing there, and if it is does this suggest with this type of measurement there is even more importance in somehow norming you're scoring key when you're administering the measure in a new culture?

DR. LEGREE: I suppose for some applications that could generally be true can you really talk about a driving culture though? Can you talk about a culture revolving around G? I'm not sure it's a sensible notion then, but surely for some applications that produces a reasonable set of questions… I'm not sure it covers all applications.

Performance and Biological Mapping of Emotional Awareness

Richard D. Lane, M.D., Ph.D.
Professor of Psychiatry, Psychology
and Neuroscience
The University of Arizona
Tucson, Arizona

PERFORMANCE AND BIOLOGICAL MAPPING OF EMOTIONAL AWARENESS

DR. ROBERTS: Our final speaker for today is Professor Richard Lane from the University of Arizona.

DR. LANE: A while ago we had the seventh inning stretch, since I'm the ninth speaker I guess that means we're in the top of the ninth. I want to thank the organizers for inviting me, and especially Rich to seeing to it that I got here. And it's really been an extremely informative experience for me and I've really enjoyed being here.

Let's just take a minute and tell you a little bit about my background and to help explain why I've done the work that I've done. I am first trained as a psychiatrist and later got a PhD. in Experimental Psychology and in my early days, I became fascinated with the question of how stress leads to and affects physical disease and that's how I got into emotion research and I became fascinated with the question trying to understand individual differences and the response to stress and how emotions were processed. And I went on to get the PhD. in part cause I'm very interested in psychometrics and interested in mind brain relationships and the brain imaging work that I've done has really been an effort to understand the physiology of emotion both centrally and peripherally and as it mediates between mental events on the one hand and physiological change and disease processes on the other.

Performance and Biological Mapping of Emotional Awareness

- Levels of Emotional Awareness

- Neural Substrates of Implicit and Explicit Emotional Processes

- Promoting emotional intelligence

Here is a basic outline on the talk, the first part major part is going to be on levels of emotional awareness which you've heard a little bit about, and I'll talk about the theory as well as the scale and measurement. The second part we'll talk about neural substrates of implicit and explicit emotional processes and I'll try to tie it a little bit to the first part as well. And if there is time, implications of the model and the findings for promoting emotional intelligence.

Emotional Awareness:
The Foundation of Emotional Intelligence

Emotional intelligence consists of a set of mental abilities:
- Perceive emotion
- Understand emotion
- Manage emotion
- Using emotion

All require conscious processing of emotional information.

Emotional intelligence: The ability to use emotional information (originating within oneself or the outside world) in a constructive and adaptive manner.

So one of the advantages of powerpoint is that you can change the slides up to the last minute and this definition was put on this slide about a half hour ago. But I think that emotional awareness is probably fundamental to emotional intelligence because we consider these different mental abilities that make up emotional intelligence, perceiving emotion, understanding emotion, managing emotion and using emotion and all do require conscious processing of emotional information.

Work on emotional awareness, I think, is kind of the foundation or substratum of emotional intelligence. One way of defending emotional intelligence from this perspective, is that it is the ability to use emotional information in a constructive and adaptive manner, with that information referring to information coming from within oneself via a kind of introspective process, as well as information coming from the outside world.
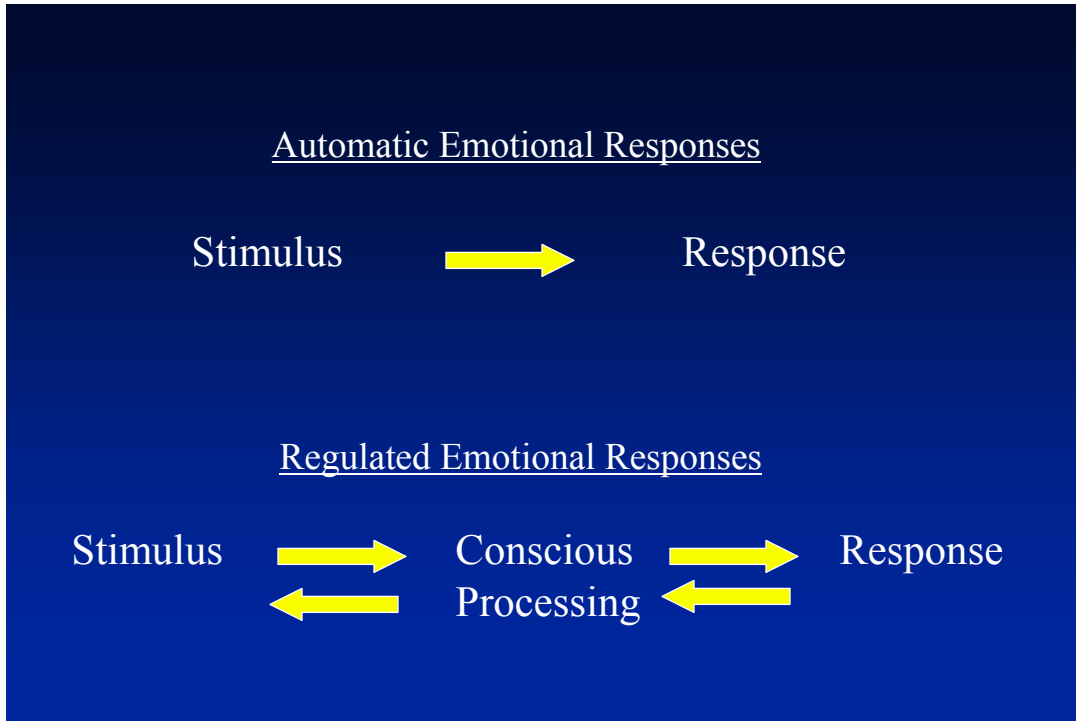
Another perspective on this is that by becoming consciously aware we have a lot that intervenes psychologically between the stimulus and response. If a car is coming at you and you have that immediate withdrawal response that is you know something that we evolved a long time ago and is a very early response, an older response and it's very automatic, and there's not much that intervenes in between these two.

But here with conscious processing if you're able to experience your emotions you can reappraise what the stimulus is before reaching a final conclusion about what the stimulus is; you can think about what your feelings are in the setting that you're in, and perhaps use that information to also think about how other people might feel if you were to respond in a particular way. So that with conscious processing really amounts to a lot of intervening cognition in between stimulus and response and in general it also corresponds to the extent of the neural network that mediates between sensory and motor processes.

THE AMERICAN JOURNAL OF PSYCHIATRY

Volume 144, Number 2   February 1987

In this issue:

Self-Portrayal by a Depressed Poet:
A Contribution to the Clinical Biography of William Cowper
By Joachim-Ernst Meyer and Ruth Meyer

Levels of Emotional Awareness:
A Cognitive-Developmental Theory and Its
Application to Psychopathology
By Richard D. Lane and Gary E. Schwartz

Official Journal of the American Psychiatric Association

So starting in on the levels of emotional awareness model I actually started working on the theory and formulated its basic outline 20 years ago. The first paper was published in February 1987 in the *American Journal of Psychiatry*. I am thankful to Gary Schwartz, my first mentor, who helped me create this model, which we called the Levels of Emotional Awareness cognitive developmental theory. In this first publication, we were also interested in its application to psychopathology.

Levels of Emotional Awareness:
Basic Principle

The structural characteristics that Piaget described for the stages of cognitive development also apply to the developmental transformations that occur in the capacity to be aware of one's own feelings.

This was a purely theoretical paper, and the basic idea was that the structural characteristics that Piaget described to stages of cognitive development also applied to the developmental transformations that occur in the capacity to be aware of ones own feelings.

The Levels of Emotional Awareness Construct

The higher the level of emotional awareness, the greater the differentiation, integration and desomatization of emotional experience and expression.

In addition, the model incorporates the idea that the higher the level of emotional awareness the greater the differentiation, integration, and desomatization of emotional experience and expression.

Wine Tasting

Compared to novices, expert wine tasters:
• Used more descriptors of wine
• Used more dimensions in their ratings
• Accurately rank ordered wines for sweetness, balance and tannin (novices accurately ranked sweetness only)

One way of understanding this perspective is to consider an analogy and that analogy is to wine tasting. A paper published in 1990 by Solomon and colleagues looked at experts and novice wine tasters, and they found that compared to novices expert wine tasters used more descriptors of wine and they used dimensions in the rating and they accurately ranked ordered wine for sweetness, balance and tannin, where novices actually ranked sweetness only, just one dimension.

You can think of that in one of two ways, one way would be people that who know more have more terms to describe the entity in this case wine and it's just a reflection of their knowledge. An alternative is that the reason they know more is that they've gone through a complex process of pairing the stimulus with verbal descriptors and over time have developed a more complex or complex schema or complex mapping of the world of wine. And so the test of that would be if you have a novice and expert sip a fine wine and ask them if the taste is the same. And how many people think that the taste would be more complex and differentiated for the experts? Okay, how many people would say the opposite, that it tastes the same?

Language, Concepts and Attention

Language facilitates the development of concepts, which influence how attention is allocated when confronted with unfamiliar stimuli.

We're adopting the view that there's an intimate relationship between language and experience. Warner and Caplan were the ones that put forward this view first; that language facilitates the development of concepts which influence how attention is allocated when confronted with unfamiliar stimuli. I'll tell you a little bit more about conceptual organization later and also come back to the question of attention when we get to some of the brain imaging findings.

Representational Redescription (Annette Karmiloff-Smith)

- The development of knowledge in any domain proceeds through a process called "representational redescription."
- Cognitive development consists of the transformation of knowledge from implicit (procedural, sensori-motor) to explicit (conscious thought) representations through the use of language or other semiotic mode.
- This transformation renders thought more flexible, adaptable and creative.
- The development of intelligence, whether cognitive or emotional, is governed by the same principles.

Now you will appreciate from the preceding slides that I put a lot of focus on Piaget and I'm sure many of you are aware that Piaget has come under some criticism and I think it is also important to point out that modern conceptions of cognitive development are consistent with the point of view that I'm putting forward and an example of that is model of Anette Karmiloff-Smith who is a protégé of Piaget.

Her view is that the development of knowledge in any domain proceeds through a process called representational re-description. Cognitive development consists of the transformation of knowledge from implicit procedural and sensory motor to explicit conscious thought representations through the use of language or other semiotic mode. This transformation renders thoughts more flexible, adaptable and creative so the development of intelligence whether cognitive or emotional governed by the same principles.

Levels of Emotional Awareness

1: bodily sensations

2: action tendencies

3: unidimensional emotional experience

4: multidimensional emotional experience (blends)

5: multidimensional experience of self and other

So we describe 5 levels of emotional awareness patterned after Piaget's stages of cognitive development. But we consider levels 1 and 2 to be implicit and levels 3 through 5 to be explicit. Level one is bodily sensations, level 2 action tendencies, level 3 through 5 involve feelings, level 3 would be a unidimensional emotional experience such as the feeling of happiness or sadness, a level 4 experience would be multidimensional emotional experience, a blended experience like feeling angry and frustrated, and level 5 is multidimensional experience of self and other where you have complexity of blends in self and other that are different from each other.
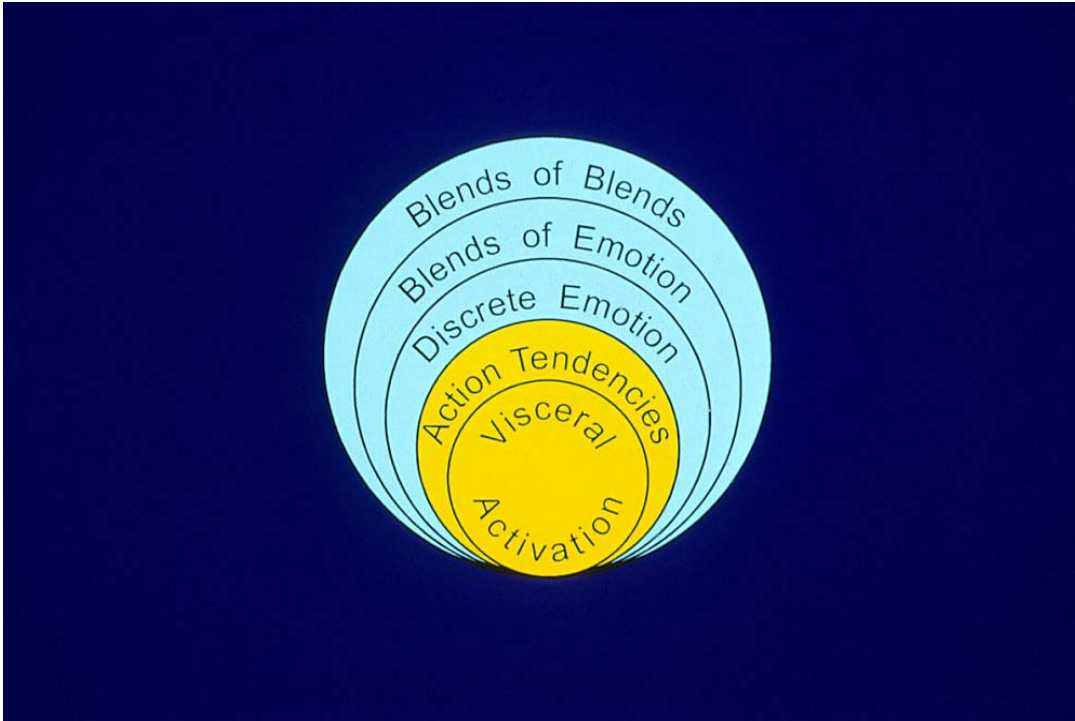
I'll show you an example of this, this is a schematic diagram of the levels intended to illustrate several points, one is what you see in yellow is implicit and what I mean by implicit is that these are visual motor and sensorimotor responses, but they are part of an emotional response, but they only could truly be considered an emotion if they're associated with a feeling. And we'll talk more about manifestations of emotion that are not such as a patient who is preoccupied with somatic sensations.

Conceptual Organization of Emotion

- Each emotion becomes a conceptual entity in its own right:
  - What the experience is like; how it feels in the body
  - What the outward signs of the emotion are
  - What causes the feeling
  - What enhances or diminishes the feeling
  - How the emotional state is related to overt behavior
  - How social context influences what is or should be expressed
- Over time, more features become part of each concept
- A network of inter-related schemata develops

So it is a hierarchical model and as you go up the levels each new level subsumes the previous ones. The model can be used either in a trait or state conceptualization. We have data supporting the trait model; I should tell you that the state interpretation hasn't yet been adequately tested; the idea is any given emotional experience would be a micro genetic construction going up the levels to the highest level that the person is at, at the moment. So the experience of happiness would include visual activation and action tendency of approach and unidimensional feeling and the fact that it is consciously experienced would lead to a different kind of visual and sensorimotor activation than if it hadn't been consciously experienced. Then a trait interpretation is the typical level that the person is functioning at.

Conceptual Organization of Emotion

The conceptual organization of emotion determines:

• The complexity of subjective experience
• The ability to recognize emotions
• The range of emotional expressions
• The extent of one's emotion vocabulary
• The capacity for empathy
• The capacity for adaptive social behavior
• The capacity for emotion self-regulation

A little bit more theory, and then we'll get into measurement issues. A few years ago I established a collaborative relationship with a woman from Geneva, Switzerland who got her Ph.D. with the Piaget group, and she thought that the initial theoretical papers didn't take it far enough, so we've written a chapter that's in the handout that talks about this extension of a theory, which is we're talking about the conceptual organization of emotion, which emotional experience, emotional awareness is a part.

So that each emotion becomes a conceptual entity in it's own right consisting what the experience is like and how it feels in the body, what the outward sign of the emotion are, what causes the feeling, what enhances or diminishes the feeling, how the emotional state is related to a work behavior, how social context influences what is or should be expressed. So it's everything you know about an emotion.

Over time more features become part of each concept and a network of interrelated schema developed. And so the theory says that the conceptual organization of emotion determines a number of things, the complexity of subjective experience, which is what the levels of emotional awareness model is about. The ability to recognize emotions, the range of emotional expressions, the extent of ones emotion vocabulary, the capacity for empathy, the capacity for depth of social behavior and the capacity for emotional social relations.

Levels of Emotional Awareness Scale

- A paper and pencil performance measure

- 20 emotion-evoking scenes, each described in 2-4 sentences, each involving two people

- How would you feel? How would the other person feel?

- Subjects can write as much or as little as needed to answer the two questions

- Each of the 20 scenes is scored separately on a 0-5 scale

All right! Well how much of this is true? We have a scale called Levels of Emotional Awareness Scale to begin to look at this. And I do want to point out that you know we really haven't changed the scale much since we first thought it up 20 years ago, and I think it's in some ways remarkable that it's worked reasonably well, but I don't think it should be considered a finished product. And I think there are other ways of kind of operationalizing the theory even with the data that we've already collected.

So it's a paper and pencil performance measure consisting of 20 emotion-evoking scenes, each described in two to four sentences, each involving two people. For each scene, two questions are asked. How would you feel? And how would the other person feel? Subjects can write as much or as little as needed to answer to the questions and each of the 20 scenes is scored separately on a zero to five-point scale. Thus, you have a maximum score of 100 with the 20-item version. So the reliability is very high, as you'll see, and it involves picking out terms that might be emotional or are emotional and using the glossary to see at what level they're at. So it's all theory driven.

Examples of Emotion Words at Different Levels

0: Alone, disbelief, puzzled, skeptical, unsure

1: Dizzy, exhausted, hot, sick, worn out

2: Good, bad, upset, rattled, stressed, excellent

3: Afraid, disgust, happy, love, shame

Level 1 terms are somatic, so they say, "I feel dizzy. I feel exhausted. I feel sick." These are Level 1 terms. Level 2 corresponds to terms that are used more commonly in a non-emotional setting, but are commonly used, often used to refer to emotions such as, "I'd feel good," and "I'd feel bad." Then Level 3 discrete emotion terms, unequivocally emotional.

Scoring of the Levels of Emotional Awareness Scale

0: Non-emotion terms

1: Bodily sensations

2: Action tendencies or non-specific terms

3: One specific emotion term

4: Two or more specific emotion terms (self or other)

5: Both self and other are at level 4 and non-identical

Since we have to account for every word, we come up with the convention that words that are cognitive, and aren't necessarily emotional, are given a score of zero, so using the term, "I'd feel puzzled" to be the equivalent of a thought. So scoring for each individual item then goes on to Level 4 if there are two or more specific emotion terms applied to the self or the other, and Level 5 if both self and other at Level 4 and they are non-identical.

So here's an example: Scene 20 from the Levels of Emotional Awareness Scale. You and your best friend are in the same line of work. There's a prize given annually to the best performance of the year. The two of you work hard to win the prize. One night the winner is announced – your friend! How would you feel and how would your friend feel?

Here's an actual Level 0 response: "I don't work hard to win prizes. My friend would probably feel that the judges knew what they were doing." A Level 1 response: "I'd feel sick about it. It's hard for me to say what my friend would feel. I would all depend on what our relationship was like and what the prize meant to her."

Examples of Responses on the LEAS

- Level 2: I'd probably feel bad about it for a few days, and try to figure out what went wrong. I'm sure my friend would be feeling really good.
- Level 3: We would both feel happy. Hey, you can't win 'em all.

Level 2: "I'd probably feel bad about it for a few days and try to figure out what went wrong. I'm sure my friend would be feeling really good." So bad and good are Level 2 for both self and other. At Level 3: "We would both feel happy. Hey, you can't win them all."

Okay, Level 4: "I would feel depressed. A friend in this light is just like any other competitor. I would also begrudgingly feel happy for my friend and rationalize that the judges had erred," so the self is at Level 4 because of the depressed and happy. "My friend would feel very gratified, but would take the prize in stride to save the relationship." So the only real emotion term there is gratified, so the other is at Level 3. Here's an example of Level 5: "I'd feel disappointed that I didn't win, but glad that if someone else did that person was my friend. My friend probably deserved it." So disappointed and glad, you're at Level 4. "My friend would feel happy and proud, but slightly worried that my feelings might be hurt."

So the LEAS has strong psychometric characteristics. It's really very highly reliable. The internal consistency as you see here is in the high 80s and the 10-item version, which is just splitting the 20-item version into two halves, also has very decent internal consistency. Inter-rater reliability is extremely high because there is essentially no inference involved. And test-retest reliability in two or three weeks .6 in the Spearman Brown is okay. And at 10 months in a small sample it was .39, and we don't know what that means. It may be because the LEAS is sensitive to state effects.

There have been a number of findings, both from my own group as well as from other people, you know well over 100 people have requested and are using the scale, but I just want to briefly summarize some of the findings. So there is a positive correlation with verbal ability. It ranges from .17 to .38.

There was a recent paper published in *The Journal of Individual Differences* and *Personality and Individual Differences* and the correlation there with the measure of vocabulary was .27 in a sample of about 100 people. Correlates with other cognitive developmental measures, such as Levenger's Sentence Completion Test of Ego Development: the correlations are kind of in the mid-20s; with Blatt's(?) Measure of the Cognitive Complexity of the Description of Parents, about the same. LEAS correlates with openness to experience in several studies. Emotion recognition accuracy: if you score higher on the LEAS you really are more accurate in emotion recognition.

We did a study of 380 individuals in a community-based sample equally stratified for the two sexes, five age groups and three socioeconomic classes, and the correlation with emotion recognition accuracy with different kinds of tasks was about .43. It correlates positively with impulse control. In practically every single study that's been done women score higher than men, even when you control for verbal ability and that's work that I've done with Lisa Feldman Barrett. Discriminate validity, it consistently does not correlate with self-reported negative affect, and also does not correlate with affect intensity, so these are two major dimensions of emotion – valence and arousal, it doesn't correlate with them, so it's doing something else.

LEAS: Recent Psychometric Findings

Positive correlation between LEAS and Empathy
(standard self-report measure)
• Positive correlations between LEAS and MSCEIT
subscales A (faces), C (blends), D (progression)
• Predicts likelihood of seeking help for emotional
problems
• Correlates with actual amount of social support
• Judgements of general well-being not influenced by
transient mood induction in high LEAS subjects

Some more recent findings include positive correlations with self-report measures of empathy. Now its correlation with the MSCEIT: Kim Barchard included the LEAS and the MSCEIT in her dissertation, about 300 people, and there were low level correlations that were significant with these subtests with the MSCEIT about .18. LEAS also predicted the likelihood of seeking help for emotional problems and correlates with actual amount of social support that people have. And in this experimental study that was just published in *Personality and Individual Differences*, they did an experimental study where they induced either positive or negative mood the between-subject study, and they had people rate their general sense of well-being before the mood induction and then after. What they found was that people who were low in emotional awareness changed their general rating of emotional satisfaction based on the mood that they were in, whereas the higher emotional awareness folks did not change.

LEAS: Clinical Findings

- Lower scores in Borderline Personality Disorder (Levine et al. 1997)

- Inverse correlation with non-specific somatic symptoms
  in fibromyalgia

- Inverse correlation with pain severity in patients with
  irritable bowel syndrome

- Inverse correlation with intrusive thoughts in patients
  with prostate cancer 6 months after Pennebaker writing exercise

- Psychosomatic Inpatient Ward - Cologne

And then there are some clinical findings. Lowest scores on the LEAS are for people with borderline personality disorder and there is an inverse correlation with nonspecific somatic symptoms and fibromyalgia. An investigator in Syracuse told me this week that in a sample of 50 Irritable Bowel Syndrome patients that their pain severity was inversely correlated with the LEAS. A recent finding at a meeting in Holland is that there's an inverse correlation with intrusive thoughts in patients with prostrate cancer six months after a pen-and-paper writing exercise-- the more emotional awareness expressed in their writing, the better they did in terms of PTSD symptoms.

German Psychosomatic Ward Inpatient Study

Caudia Subic-Wrana, Ph.D

Susanne Bruder, Ph.D.

Walter Thomas, Ph.D.

Karl Köhle, M.D.

And then I want to briefly tell you about an inpatient study on a German psychosomatic ward where they have patients, lots of patients with somatoform disorders. This involved consecutive inpatients admitted to a ward, well two different wards of two general hospitals over a three-year period in Cologne, Germany.

German Psychosomatic Ward Inpatient Study

- Consecutive inpatients admitted to the Psychosomatic Wards of 2 general hospitals over a 3-year period in Cologne
- Diagnostic groups: 1) depression; 2) anxiety and OCD; 3) adjustment disorders; 4) somatoform disorders; 5) psychological factors with somatic disorders; 6) eating disorders; 249 patients completed treatment
- TAS-20, LEAS (10-item), STAI, SCL-90 administered at onset and end of treatment (2-3 months)
- Multi-modal treatment: psychodynamically-oriented individual and group psychotherapy, body-related therapies, art therapy, music therapy, and medication as needed

There were six different diagnostic groups. They admitted those with depression, anxiety and OCD, adjustment disorders, somatoform disorders, psychological factors associated with a somatic condition, and then eating disorders. There were 249 patients who completed treatment. The main point of this is that there are advantages of LEAS relative to the Toronto Alexithymia Scale in a clinical setting. The TAS-20, as you know, is a self-report measure. This is, the LEAS is more difficult to administer and score. Is there a pay-off? I think these data will show you that there is. Patients had measures of negative anxiety – I'm sorry, negative affect, anxiety, and SCL-90. And the patients received multimodal treatment, which was a psycho-dynamically oriented individual and group psychotherapy, body-related therapies, art therapy and music therapy and medication as needed. The TAS-20 correlated highly, or significantly, with measures of negative affect both at onset and at the end of treatment.

Correlations with Negative Affect

	Onset (n=394)		End (n=249)	
	TAS-20	LEAS	TAS-20	LEAS
SCL-90	.36**	.06	.37**	.08
STAI	.34**	.15*	.42**	-.10

LEAS generally did not correlate with negative affect. Across the six diagnostic groups there were no differences between the somatoform patients and the other patients with mental disorders in TAS-20 at onset of treatment.

Change in TAS-20 With Treatment

TAS-20	DEP n=54	ANX n=20	ADJ n=31	SOM n=68	PFS n=38	ED n=38
Onset	56.67 (11.8)	54.41 (13.0)	55.03 (12.1)	52.69 (11.9)	54.97 (11.4)	55.64 (10.3)
End	50.69 (11.2)	51.16 (12.2)	52.18 (11.1)	50.98 (12.2)	51.45 (11.2)	51.30 (11.4)
	$p<.001$	NS	NS	NS	$p<.05$	NS

In all groups the TAS-20 score went down. In other words, it appeared that alexithymia improved. However, when you adjusted for negative affect, these changes were not significant.

Change in LEAS with Treatment

Treatment Completers (n=249)

LEAS	DEP n=54	ANX n=20	ADJ n=31	SOM n=68	PFS n=38	ED n=38
Onset	27.29 (4.6)	28.9 (3.9)	29.47 (6.4)	25.86 (6.5)	26.24 (5.6)	31.86 (5.2)
End	27.66 (7.3)	28.01 (5.2)	29.96 (4.7)	27.7 (7.1)	28.33 (6.5)	32.57 (7.0)
	NS	NS	NS	p<.05	p<.05	NS

With the LEAS there was a difference at onset of treatment where the somatoform patients and those with psychological factors affecting physical condition had lower LEAS scores than in the patients with mental disorders, and this is using the 10-item version. And then within the somatoform groups, there was a significant increase from beginning to end of treatment, and when you partialed out negative affect, these findings were still significant.

Neural Substrates of Implicit and Explicit Emotional Processing

• Reflective awareness

• Focal Attention to Feelings

• Background Feelings

⎫
⎬   Explicit
⎭

• Implicit Emotion

So I think now we're ready for the bottom of the ninth. Because of time limitations, I'm not going to be able to give as thorough an explanation, present as many findings to support the model that I'm going to present, but the basic idea is that we can distinguish between neural substrates of implicit and explicit emotional processes. And I think that the evidence is there to support the idea that there are three different aspects of conscious awareness of emotion that can be distinguished, that have distinct neural substrates, and those are what I'm calling reflective awareness, focal attention to feelings and background feelings. I'm going to start off with focal attention to feelings because that was a study that involved the Levels of Emotional Awareness Scale.

Explicit Emotional Processing: Highlighted Structures
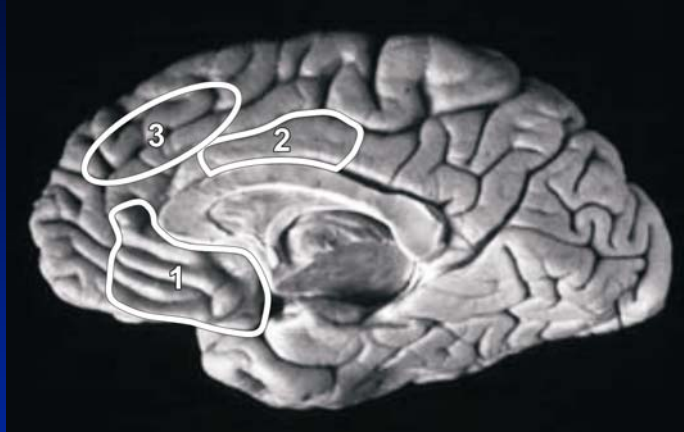
1 – Background Feelings
2 – Focal Attention to Feelings
3 - Reflective Awareness of Feelings

And just as a preview, we're going to be focusing on focal attention to feelings on this area, which is the dorsal anterior cingulate cortex, which is an area that has a super-eminent role in the regulation of attention. This area three here, superior medial prefrontal cortex, is involved in reflective awareness of feelings, and that's an area identified by Frith and Frith of being involved in mentalizing. And then we have this ventromedial prefrontal area involved in background feelings, which I'll describe for you in a few minutes.

Focal Attention to Feelings

I wish to turn now to discuss focal attention to feelings.

PET Study of Emotion

Happy
Sad
Disgust
Neutral (x3)　　　　　Film Clips

Happy
Sad
Disgust
Neutral (x3)　　　　　Recall

We did a PET study of emotion. We collected the data about 10 years ago now, and we induced happiness, sadness and disgust, as well as neutral control conditions in two different ways – with film clips and recall personal experiences. We wanted to have comparable conditions for the two kinds of ways of eliciting emotion. When we had people recall experiences of sadness, for example, we wanted them to focus in on a specific aspect of the memory where the feeling of sadness was particularly intense, like focus on that moment when the casket was being lowered into the grave, and we told them focus on feeling sad. Then we did the injection and we did a one-minute scan.

So to have it be comparable for the film clips, we told them, "Okay, this is going to be a happy film. Here's the background information." And then we showed the film clip and did the scan as they were watching the film. The point is that we were having them focus on the particular target emotion. And what we found was that when we basically lumped the three emotion conditions together, subtracted out the neutral to control for non-specific factors, so we were identifying brain activity specifically attributable to emotion, and correlated that with LEAS in these 12 healthy women, there was only one area that showed a significant correlation and that was in this dorsal anterior cingulate area.

Positive Correlation:
Emotional Awareness and Dorsal ACC Activity

Finding: Higher scores on the LEAS are associated
with greater activity in the dorsal anterior cingulate
cortex during emotional arousal.

Interpretation: The dorsal ACC has a superordinate role
in the executive control of attention and motor responses.
The dorsal ACC is activated in response to a variety of
cognitive and emotional stimuli. Those higher is emotional
awareness may be better able to attend to their own
emotional experiences, promoting further emotional development.

So the finding is higher scores on the LEAS are associated with greater activity in the dorsal anterior cingulate cortex during emotional arousal. What does that mean? Well, this is our interpretation, that the dorsal anterior cingulate has a super-eminent role in the executive control of attention and motor responses. The dorsal anterior cingulate is activated in response to a variety of cognitive and emotional stimuli, such as pain or even mental effort. The idea is that those higher in emotional awareness are activating this area because they may be better able to attent to their own emotional experiences and therefore engage attentional mechanisms when emotion is aroused. And to the extent that they do that, that will then promote further emotional development. This was published in *The Journal of Cognitive Neuroscience* in 1998. There are a variety of other findings that are consistent with that, but I'm going to have to move on in the interest of time.

Now, I would like to turn to reflective awareness and issues surrounding this concept.

Selective Attention to Pictures

- **Internal Focus: Subjective Experience**

  Pleasant   Unpleasant   Neither

- **External Focus: Spatial Location**

  Indoors   Outdoors   Either

We created another paradigm to try to capture conscious experience of emotion, and we used International Affective Picture System pictures, and we had people focus on the pictures in two different ways. In half the scans we had them focus on the subjective experience that was induced by the pictures. They had a three-key keypad and they had to press the key as soon as the picture came up. Did it induce a pleasant feeling, an unpleasant feeling or a neutral feeling? And the other half of the scans in this PET study we focused on spatial location. Was this an indoor scene, an outdoor scene or indeterminate? Now this wasn't simply focal attention because they're also labeling it with the keypad response, so that's why we think it's revealing neural substrates of reflective awareness.

So here are examples of IPS pictures. The picture on the right hand side is neutral.

Lane et al. Neuroreport 8: 3969-3972, 1997
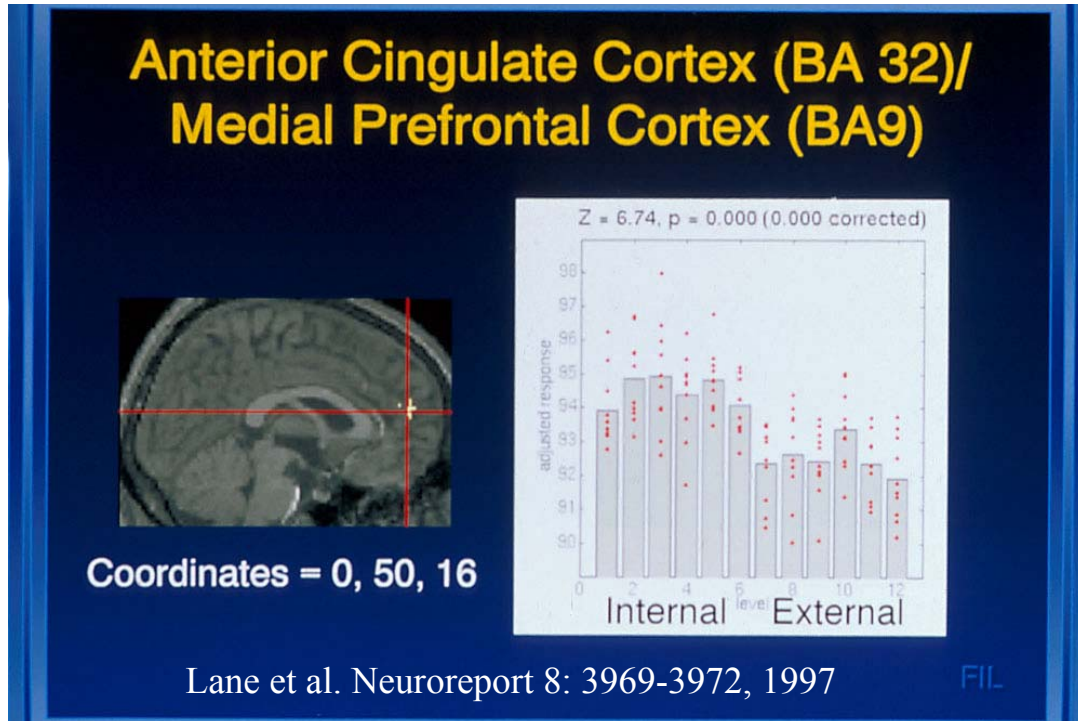
And we got very clear results and highly significant results showing really highly significant activity here in the medial prefrontal cortex in the peri-cingulate region when people were focusing on their internal experience. Well, just like the finding with emotional awareness in the dorsal anterior cingulate cortex, it's not an emotion-specific area, but rather an attention area. What we're finding here is that we've got an activation in this paradigm because we're recruiting an area involved in a more general cognitive process, that is a mentalizing.
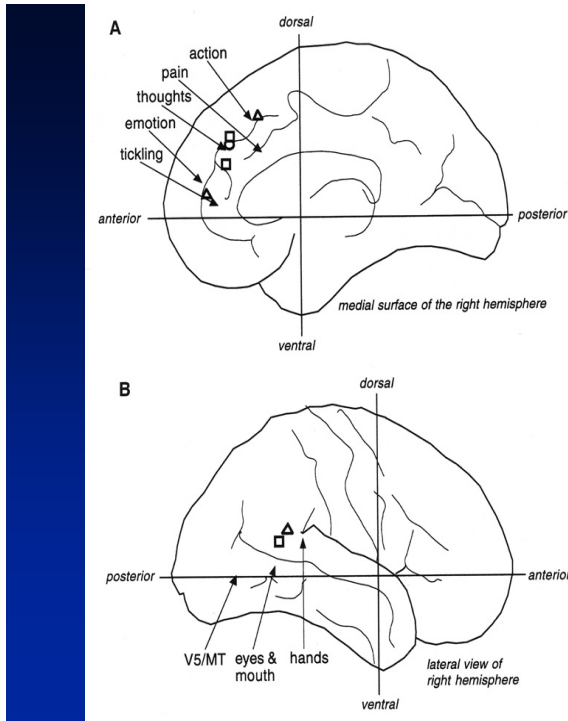
C. Frith and U. Frith

Interacting Minds -
A Biological Basis

Science 1999;
286: 1692-1695
11/26/99

Frith and Frith say that this is a skill that's developed in people and in higher apes to understand the intentions of other animals or other people. And that the ability to self monitor comes first from the ability to need to know what's going on in other people's minds, because there are evolutionary advantages to knowing that, whether you're being deceived for example.

Background Feelings

I'd like to change focus now and discuss background feelings.

Background Feelings

- Consist of the on-line experience of feelings
   in the absence of attending to them or reflecting
   upon them
- Add emotional tone to experience but are
   not noticed unless attended to
- Provide information about the general state of
   well-being – in the background of cs. experience
- Bias decision-making and behavioral response to
   the environmental "cause" of the emotion

This is a term introduced by Damasio and I have a somewhat different interpretation of it, a little shifting of the use of the term. This is how I use it. It consists of the online experience of feelings in the absence of attending to them or reflecting upon them. It adds emotional tone to experience, but background feelings are not noticed unless attended to. It provides information about the general state of well-being, but it's in the background of conscious experience. It may be related, certainly overlaps with Lisa Feldman Barrett's concept of core affect, and it biases decision making and behavioral responses to the environmental cause of the emotion.
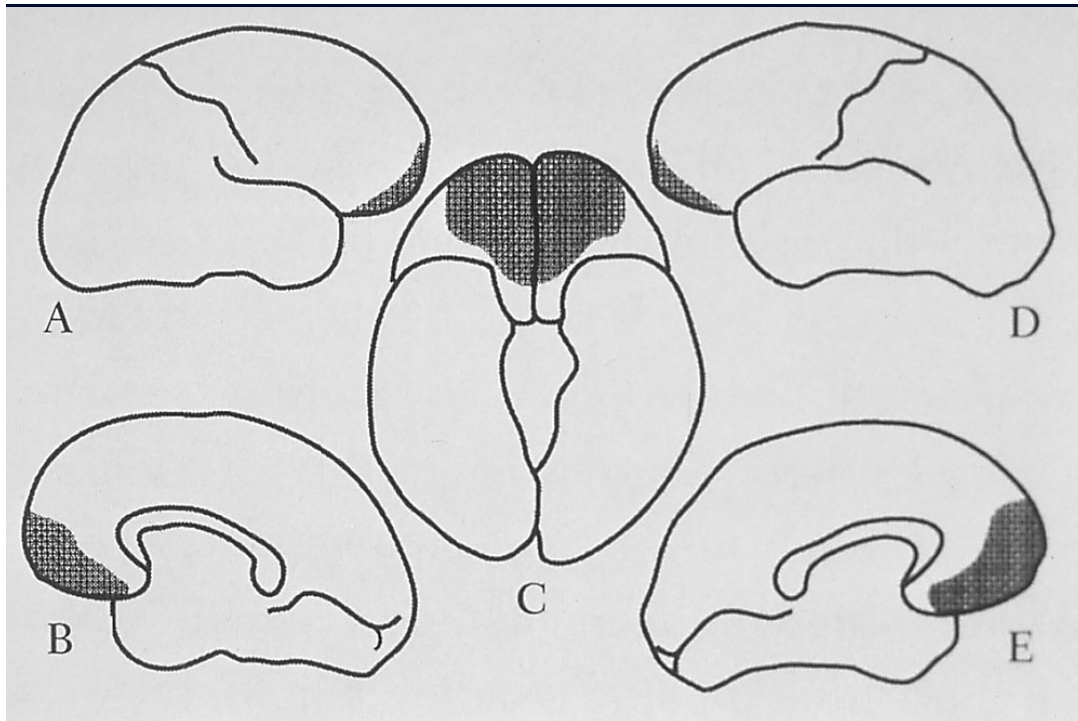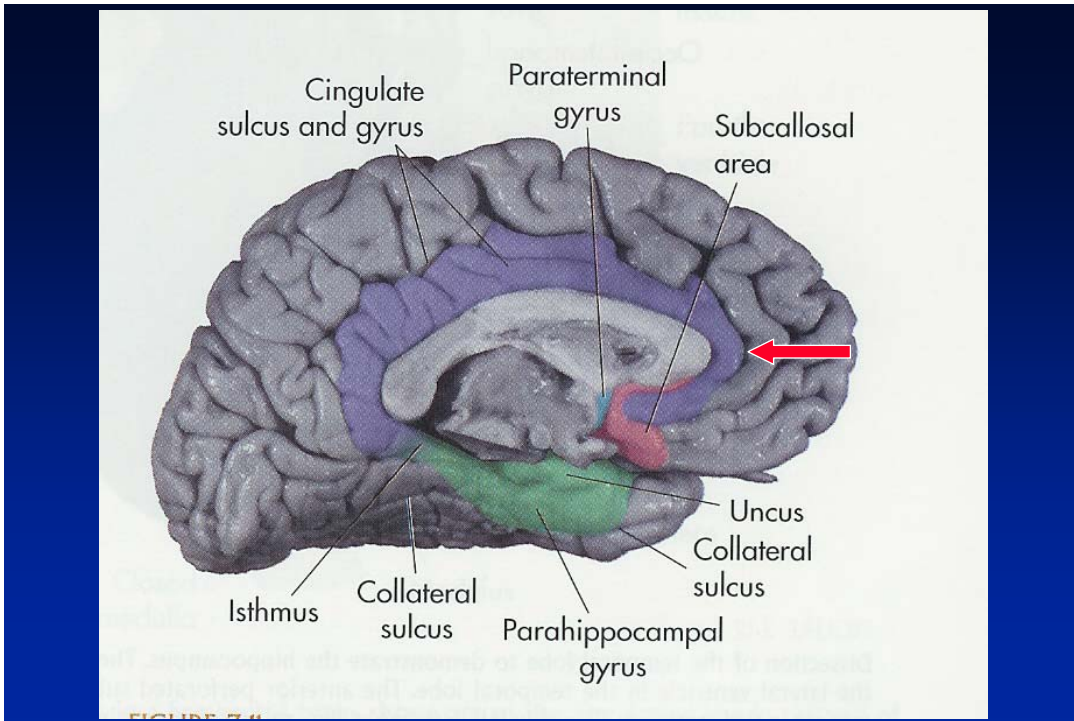
So the work by Damasio and his colleagues on ventral medial prefrontal cortex, the somatic marker hypothesis, the people who have lesions here are not able to use their gut feeling in making decisions is one example …
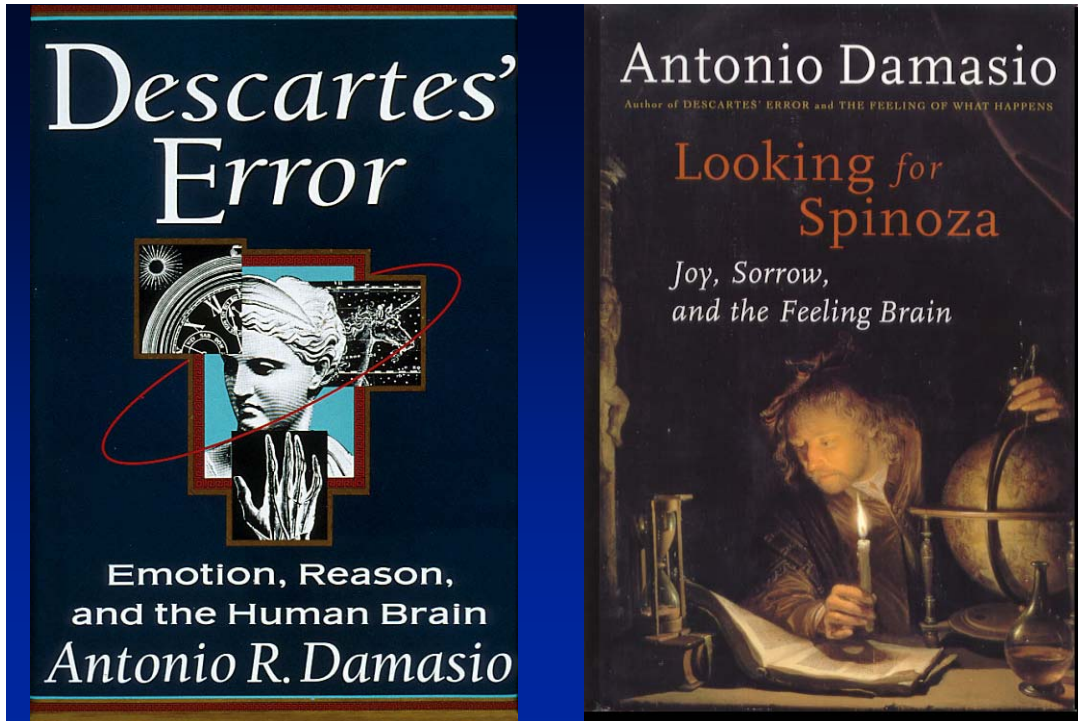
And I certainly recommend these books to you by Damasio. There are a number of other areas that are involved in background feelings.
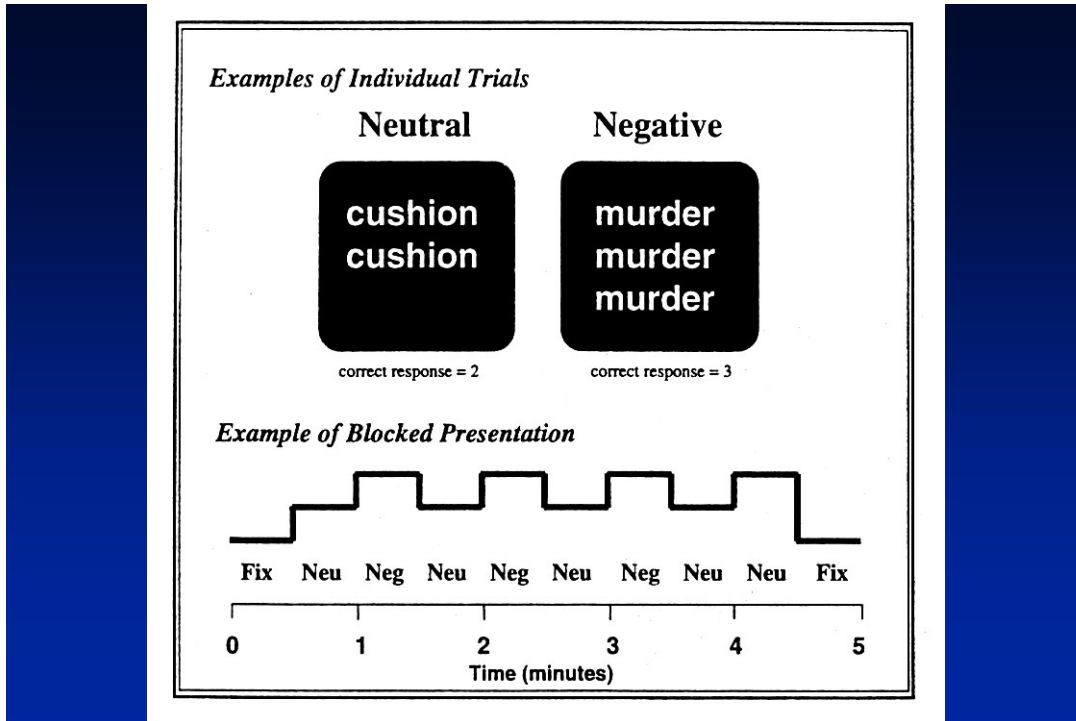
A nearby area to the one that I showed before is – we're calling pregenual inter-cingulate cortex. I'm just going to show you an example of a paradigm that activates this region.

This is the emotional counting stroop, so the task involves counting the number of words. So a neutral control condition would be words that are neutral in meaning, like cushion. The correct response is two. And this, with the word murder, would be the negative emotional condition – you have to say three. The point is that figuring out how many words there are is in the foreground. That's what you're focusing on. The emotion is more in the background.
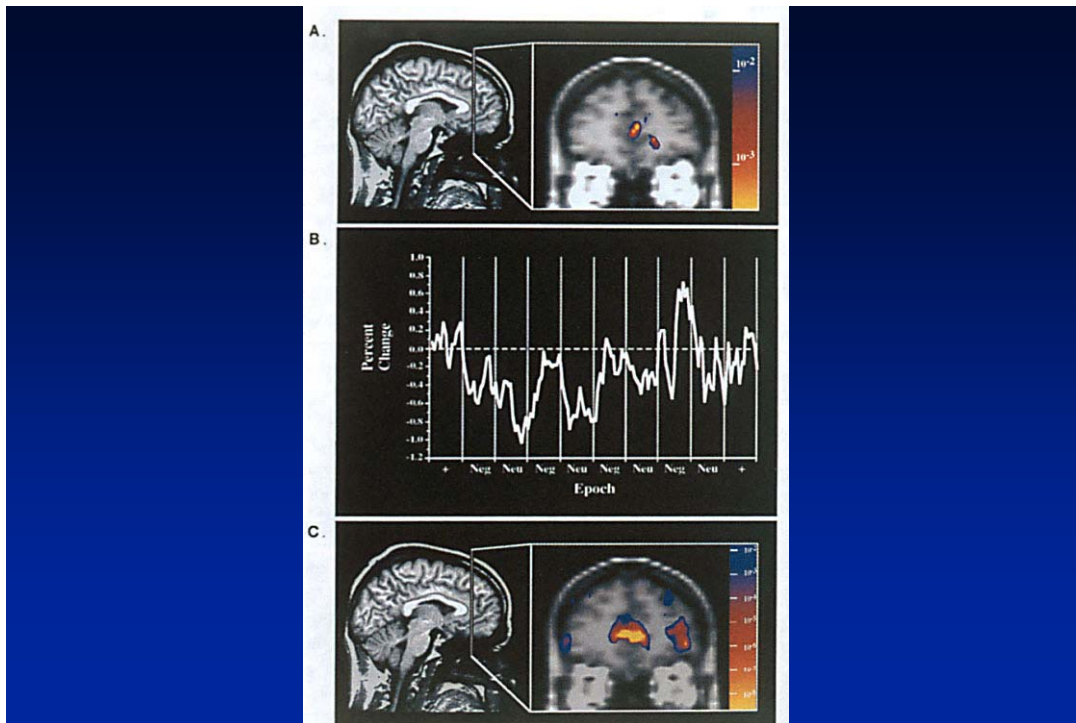
And when performing the emotional counting stroop, we're getting activity here in pregenual inter-cingulate.

Brain Structures That Participate in Background Feelings

- Ventromedial prefrontal cortex
- Pregenual anterior cingulate cortex
- Insula
- Somatosensory cortex
- Right inferior parietal cortex

These are some of the other structures involved in background feelings. For those of you who are familiar with Damasio's work, you'll see that these are areas that he talks about as being involved in feeling, and I'm saying this is background feeling, and in order to – if you actually focused your attention on your feelings, then you recruit a different area, and if then you reflect up them. You label them and otherwise process them, yet another area is involved.

Implicit Emotion

What about implicit emotion?
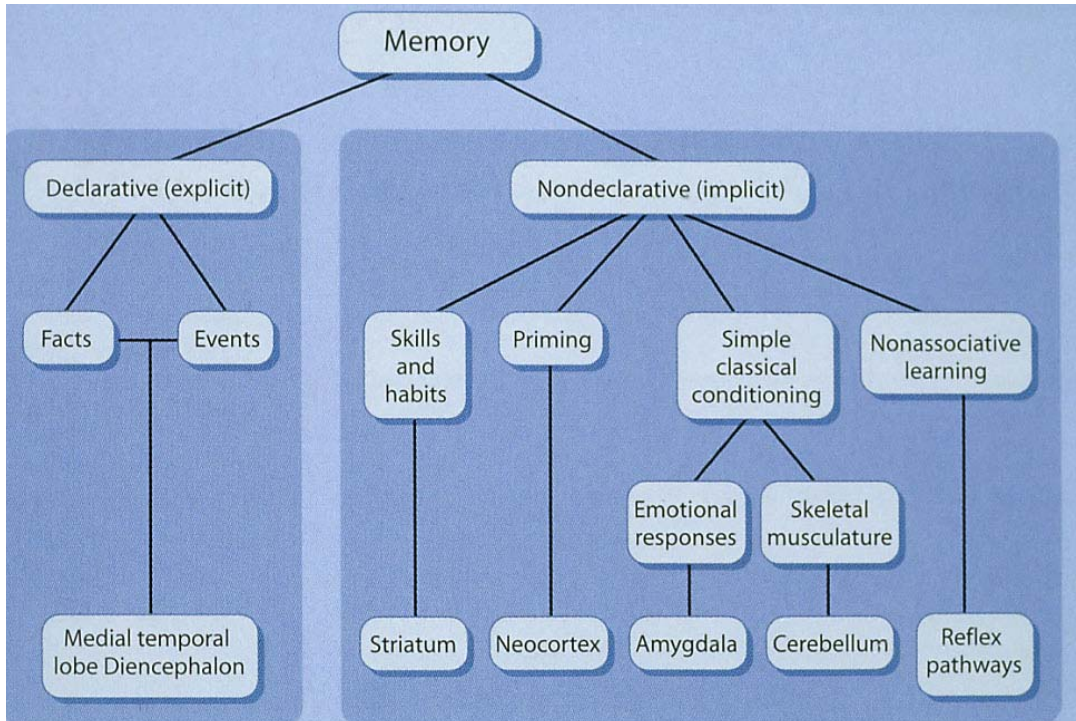
Well, the distinction between implicit and explicit processes is very fundamental in cognitive neuroscience and it was I think first best described in terms of memory, where we have declarative or explicit memory for facts and events, the ability to consciously recall these things. And this is an important distinction because the neural substrates for the two processes are distinct.
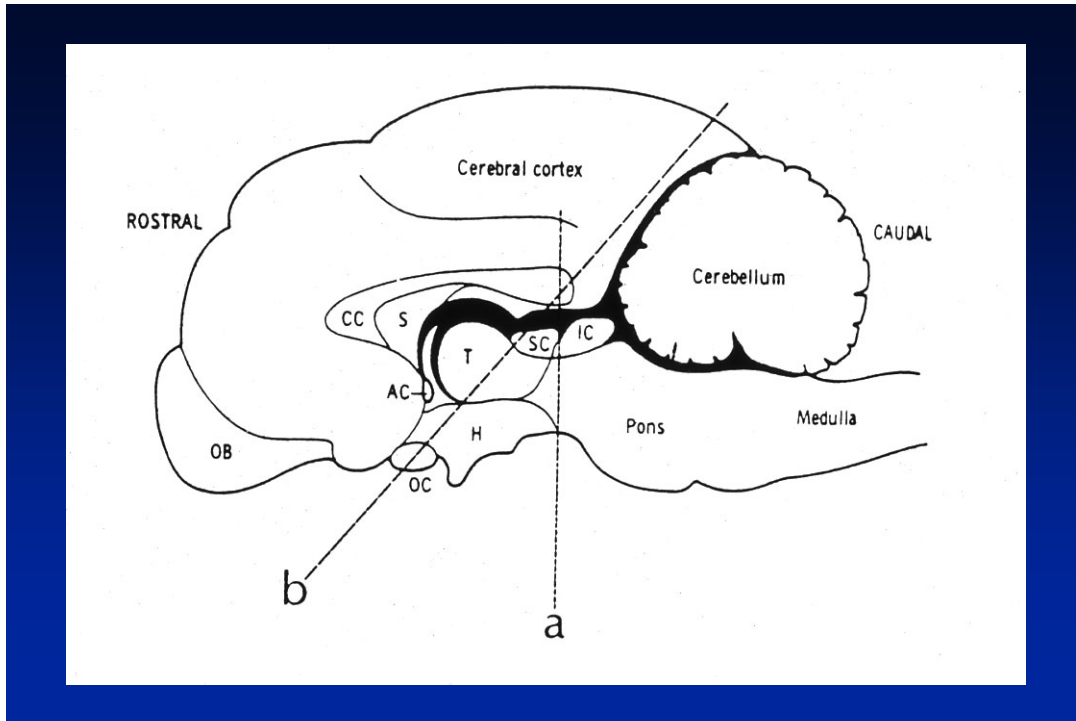
So for explicit memory we have the hippocampus, and hippocampal area and the diencephalon on the one hand. And then for implicit memory, skills and habits, priming, simple classical conditioning and non-associative learning, each is associated with a different kind of neural substrate, so that in anephric patients, for example, can be taught a task and they won't remember that they've been taught. They won't know that they know it, but then they can do it. So that kind of procedural knowledge, it's implicit memory.

And in fact, it's with the recognition that there are implicit and explicit cognitive processes that the boundaries between emotion and cognition I think has pretty much disappeared. In terms of the neural substrates of these visceral motor and somato-motor components of emotion that I'm calling implicit emotion, Cannon and Barge showed in the 1920s that if you made a lesion here at A, you eliminated all emotional reflexes and behavior, but if you made a lesion at B, keeping parts of the thalamus and hypothalamus intact, you still had emotional reflexes such as sham rage, so the display of rage in response to innocuous stimulant.
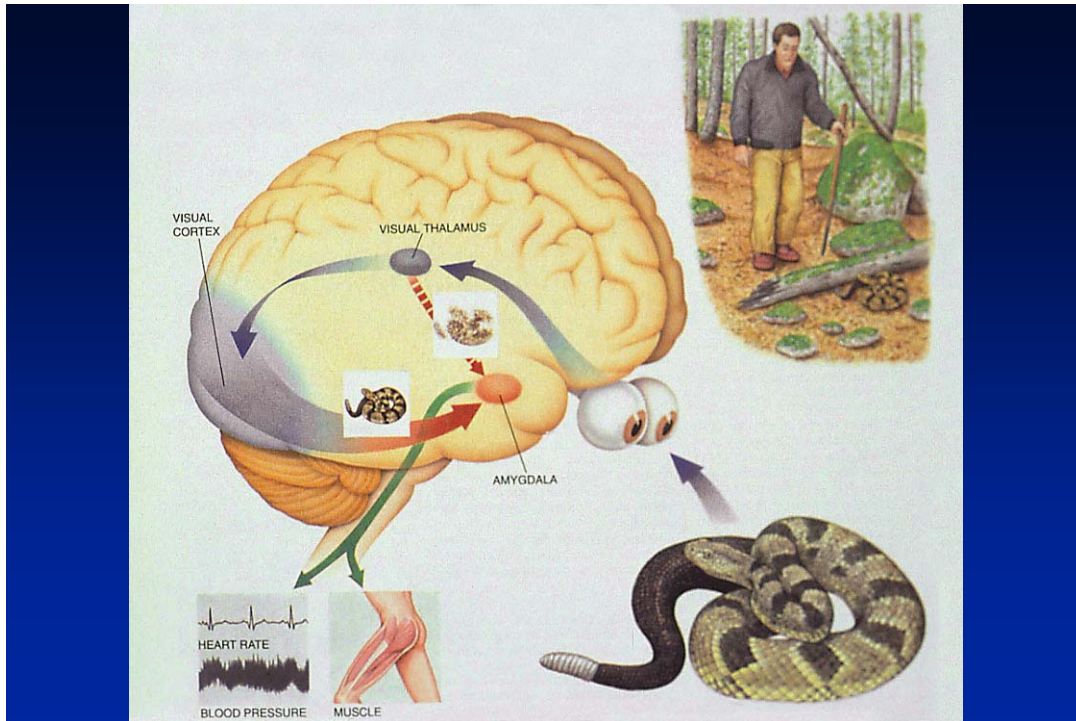
And we've heard about the low road and the high road in other talks. This slide is from Ledoux's book on the amygdala. The book is called *The Emotional Brain*, but it's mostly focused on the amygdala. And so what this shows is that this man is in the forest. He encounters what might be a coil of rope or it might be a snake. And so information goes to lateral geniculate and the thalamus. And then there is a low road pathway immediately to the amygdala, and there's this kind of crude representation. It might be a rope. It might be a snake. There's an advantage to thinking and having a response as if it might be a snake, it might be inaccurate. That happens out of awareness and it needs to happen out of awareness because you need to respond very quickly. Then there's the high road where you can see exactly what kind of snake it is, etcetera. The amygdala sends out command patterns to organize autonomic and somato-motor responses. It participates in a wider network that does that.

Left Amygdala/SI (y = 0; see Figure 3)   Right Amygdala (y = -6; see Figure 2)
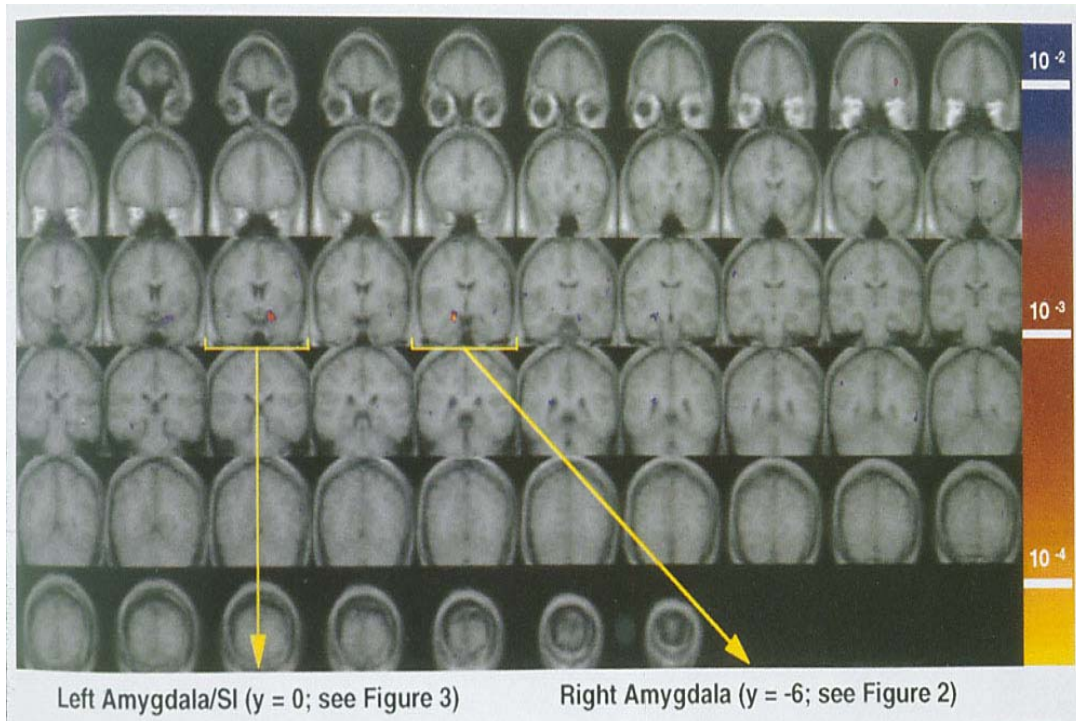
Some of the first evidence that the amygdala can operate implicitly was presented in 1998, in the *Journal of Neuroscience*. Paul Waling showed that the amygdala was activated in response to briefly-presented angry faces that were followed by a backward mask consisting of a neutral face. All people saw was the neutral face, but when the angry faces were presented very briefly subliminally, the amygdala was activated because it's a threat detector.
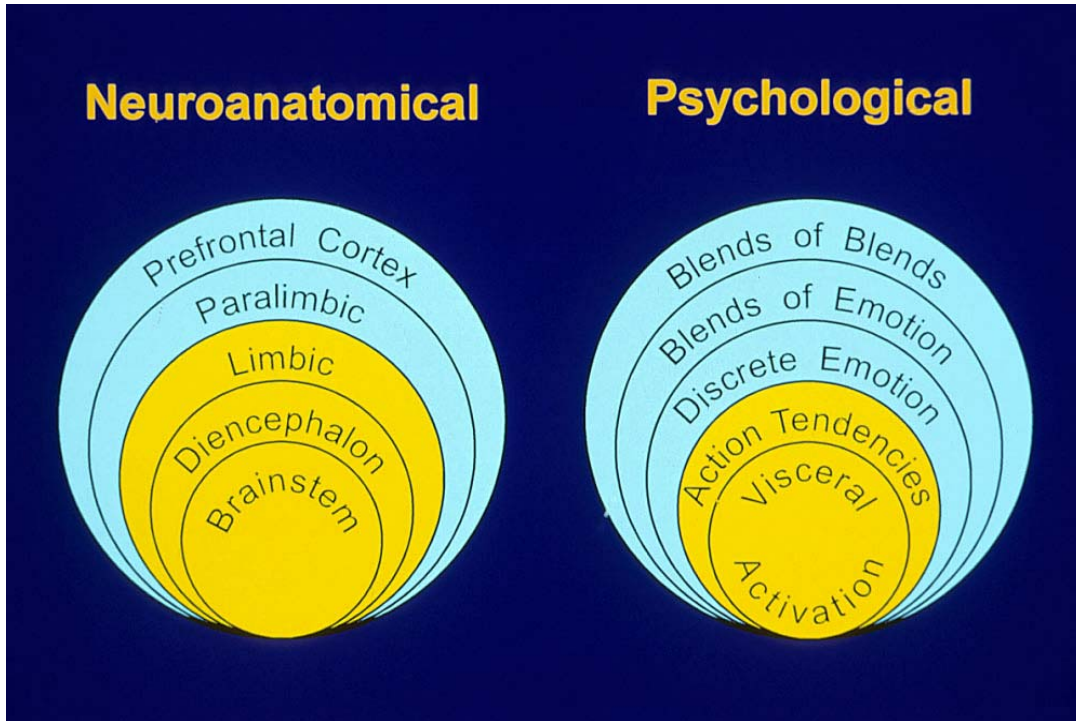
So to kind of summarize and show parallels between psychological and the neuroanatomical, you have a hierarchical neuroanatomical model here and what I'd argue is that there's broad correspondences between what you see in yellow here and what you see in yellow there. The amygdala and other limbic structures perform their functions outside of conscious awareness, and that their output has to go to higher centers to paralimbic structures such as the anterior cingulate cortex, orbital frontal cortex and anterior temporal pole as well as the prefrontal, superior medial prefrontal cortex for the conscious experience of emotion.

SERIES IN AFFECTIVE SCIENCE

COGNITIVE NEUROSCIENCE OF EMOTION

Edited by
RICHARD D. LANE
LYNN NADEL
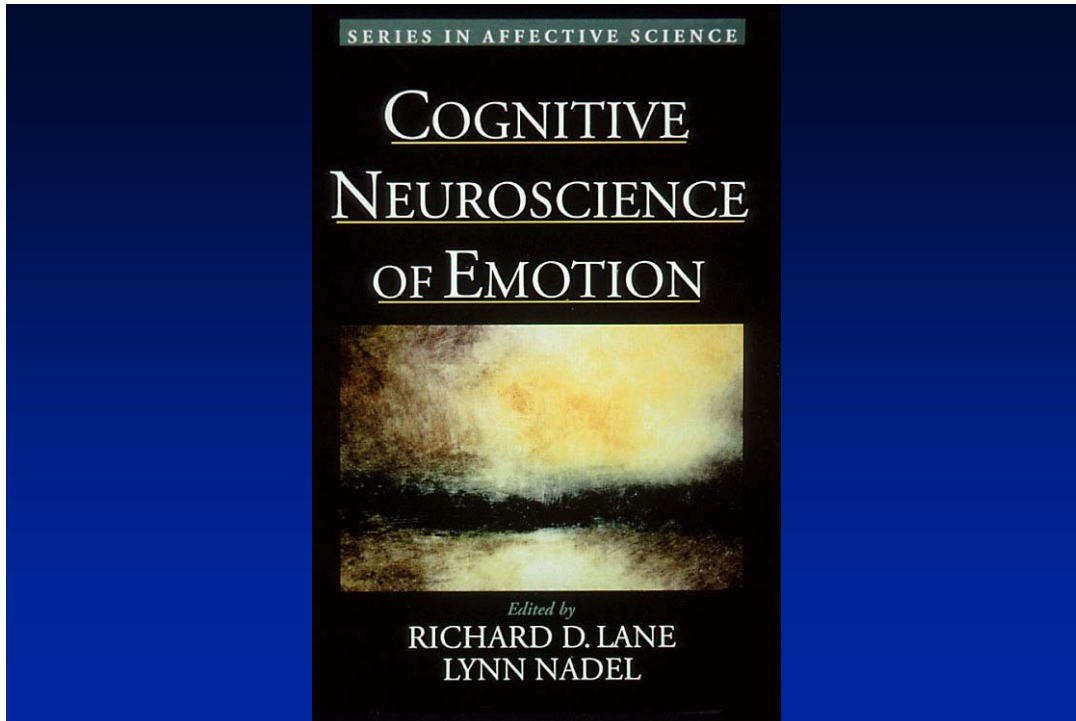
This model and the justification for it is elaborated in this book that I co-edited with Lynn Nadel. It's available in paperback. I don't get any royalties, but it's published by Oxford University Press in 2000.

Implicit Emotion and
Emotion Regulation

Implicit emotion: the automatic <u>motor</u> expressions of emotion that get modulated by emotion regulation strategies, including autonomic, neuroendocrine, and somatomotor {gestures, facial expressions, action tendencies, procedures} responses.

So implicit emotion then is the automatic motor expressions of emotion that get modulated by emotion regulation strategies, including autonomic neuroendocrine and somato-motor responses, the somato-motor responses including gestures, facial expressions, action tendencies and procedures or scripts.

Emotional Awareness and the Brain: Take Home Messages

1. The fundamental distinction between implicit and explicit processes applies to emotion as well as cognition.
2. Emotion generation occurs without awareness of feelings.
3. The neural substrates of three types of conscious experience can be identified:
    - a. Background feelings
    - b. Focal attention to feelings
    - c. Reflective awareness of feelings
4. The neural substrates of conscious emotional experience are not unique to emotion but are shared with other cognitive functions, setting the stage for individual differences.

Here are the main points regarding the neural substrates component of the talk. The fundamental distinction between implicit and explicit processes applies to emotion as well as cognition. Emotion generation occurs without awareness of feeling. The neural substrates of three types of conscious experience can be identified – background feelings, focal attention, reflective awareness – and the neural substrates of conscious emotional experience are not unique to emotion, but are shared with other cognitive functions, setting the stage for individual differences. So that's a particularly important point.

Promoting Emotional Intelligence

- In normative contexts

- In contexts of maladaptive behavior

What are the implications of this model for promoting emotional intelligence? You can divide it into normative context and then context of maladaptive behavior.

Conceptual Organization of Emotion: A Psycho-Educational Approach

- Each emotion becomes a conceptual entity in its own right:
  - What the experience is like; how it feels in the body
  - What the outward signs of the emotion are
  - What causes the feeling
  - What enhances or diminishes the feeling
  - How the emotional state is related to overt behavior
  - How social context influences what is or should be expressed
- Over time, more features become part of each concept
- A network of inter-related schemata develops

It occurred to me that this conceptual organization of emotion could be a guide to promoting emotional awareness in a kind of psycho-educational approach, so teaching of people about emotions individually – what the experience is like, how it feels in the body, what the outward signs are, etcetera. I don't know about the extent to which such didactic programs have already been created, but perhaps it might work. And I think it's important, as a psychiatrist I have a sensitivity to the fact that psycho-educational approaches will work in quite a few people, but then there are some people who are going to be resistant to change and aren't going to be able to pick it up that readily.

Changing Behavior Patterns
to Improve Emotional Intelligence

- Learning to drive a stick shift: with practice, automatic procedures eliminate the need for conscious processing

- Maladaptive behavior patterns that people want or need to change (e.g. the "insensitive manager") can be conceptualized as automatic emotional procedures

And so I think that it's useful to think in terms of procedures and how maladaptive behavior patterns that people want or need to change can be conceptualized as automatic emotional procedures. An example was given yesterday of the insensitive manager. And so that's a general behavior pattern. Well how do you go about changing that? I think that it's important to point out that in terms of thinking about procedures, learning to drive a stick shift that it first requires a lot of conscious processing, then it becomes automatic. And it's useful for these things to become automatic because it eliminates the need for conscious processing. It's very economical. Well how do you undo it? Well I think that you have to go from implicit to explicit processing.

Emotional Procedures

Rule-based schemas for <u>how</u> <u>to</u>:

- Express love
- Obtain love and reassurance
- Handle anger
- Get attention
- Joke around
- Resolve conflict

Emotional procedures are rule-based schemas for knowing how to do things related to emotion – express love, obtain love and reassurance, handle anger, get attention, joke around, resolve conflict, etc. This kind of approach has been written about in the clinical literature and I'll try to tell you what this guy Robert Klinenman said in 1991.

Procedural Organization of Emotions

Clyman R. B. J Am Psychoanal Assn 39: 349-382, 1991

- During emotional arousal component emotional procedures (facial, gestural, overt behavioral) are activated

- Emotional procedures are repeated in different contexts

- Families "select" which procedures are permissable and which are not

- Procedures reflect how self and other are expected to behave

During emotional arousal, component emotional procedures are activated, these procedures including facial, gestural, overt behavioral expressions of emotion. These emotional procedures are repeated in different contexts and we select which procedures are permissible and which are not. Procedures reflect how self and other are expected to behave. These procedures develop implicitly and with practice are executed automatically.

Changing Emotional Procedures

- Emotional procedures develop implicitly and, with practice, are executed automatically
- Once established, procedures are resistant to change
- Procedures preserve cognitive resources for conscious processing
- Maladaptive behavior patterns can be conceptualized as emotional procedures
- In circumstances in which simple instruction in alternative procedures is insufficient, changing the behavior requires first making the implicit emotions associated with the procedures explicit (controlled), then developing and practicing new procedures
- Making new procedures automatic can take a lot of practice

Once established, they're resistant to change. Procedures preserve cognitive resources for conscious processing. And we can understand maladaptive behavior patterns in this way. In circumstances in which simple instructions and alternative procedures are insufficient, changing the behavior requires first making the implicit emotions associated with the procedures explicit and controlled, and probably in order to do that you have to stop the behavior first and ask, interrogate what the feelings are associated with that action tendency. And then developing and practicing new procedures and making new procedures automatic can take a lot of practice. And I'm a golfer, and one of the rules of thumb that I've heard is that if you want to change your golf swing it takes about 2,000 practice swings to get it automatic.

Conclusions

- The skills that constitute emotional intelligence require conscious processing of feelings and exteroceptive emotion cues
- A cognitive-developmental approach can be used to understand individual differences in the capacity to experience emotion and recognize emotional states in others
- Cognitive development consists of the transformation of implicit into explicit representations
- This transformation has a neuroanatomical basis

Okay, so conclusions.

Conclusions

- Implicit and explicit emotional processes have distinct neural substrates
- The neural substrates of conscious emotional experience are not unique to emotion but are shared with other cognitive functions
- This helps to explain the need for emotional procedures, and the existence of vast individual differences in emotional intelligence
- Promoting emotional intelligence may be facilitated by focusing on shifts from one level to the next within a cognitive-developmental framework

The skills that constitute emotional intelligence require conscious processing of feelings and exteroceptive emotion cues. A cognitive development approach can be used to understand individual differences in the capacity to experience emotion and recognize emotional states in others. Cognitive development consists of the transformation of implicit into explicit representations, and this transformation I think has a neuroanatomical basis, because implicit and explicit emotional processes have distinct neural substrates. The neural substrates of conscious emotional experience are not unique to emotion, but are shared with other cognitive functions. This helps to explain the need for emotional procedures as well as the existence of vast individual differences in emotional intelligence. Promoting emotional intelligence may be facilitated by focusing on shifts from one level to the next within a cognitive developmental framework. Thank you.

DISCUSSION

DISCUSSION

DR. GADE: Maybe this is a naïve, probably is a naïve question. Are you saying that emotional responses are just like any other automatic response except they're different by the, you can differentiate them by where they occur in the brain?

DR. LANE: Just like any other automatic response?

DR. GADE: Once you've got an automatic response in place, you're not thinking about it or feeling it, like you were talking about driving a stick shift car and doing that automatically, and stopping for a traffic light, for example. I wouldn't call that an emotional response. I'd call that a learned automatic response. And you differentiate these based on the areas of the brain that get activated? Is that it?

DR. LANE: Yeah. I think that's what I'm saying, that I think that children learn how to negotiate their interpersonal world by learning how to do things: How to interact with people, how to express affection, how to resolve conflict, etcetera. And I think once a particular interpersonal situation is construed in a particular way, that automatic procedures do get set off. Is that; what do you think about that?

DR. GADE: One is more of a learned response and the other one you're saying is automatically wired in.

DR. LANE: Oh, no, no. No, I'm not saying it's wired in.

DR. GADE: Emotional response is not wired in? You think that's [cross-talk]?

DR. LANE: Let me clarify. The emotional response is kind of part of our genetic endowment. But I think that these procedures have an emotion regulatory function and they kind of keep emotions in check. It's a way of managing emotions.

DR. GADE: But you don't; you've already started the behavior in many ways and that's what you're saying. We've already started the emotional behavior before you become aware that this action is taking place.

DR. LANE: That's right.

PARTICIPANT: And the difficult part is how do you then delay this kind of response on a neurological basis? Do you see that kind of thing happening?

DR. LANE: Yeah, well I mean I think Paul Ekman's discussion of these Buddhist monks is very compelling, and I think what it; I mean we do know that the frontal lobes have direct projects to these limbic and brain stem structures, so I would be willing to bet that that's how it works, that they're just very highly sensitive to interceptive cues and have massive downward inhibition to regulate the evolution of the emotional response.

PARTICIPANT: I couldn't follow a lot of your discussion because neuropsychology is not my forte. But I was wondering if there is any research on people who are dwelling too much on their emotions and how that impacts them. Because I can see the benefits of focusing on your emotions, but it also seems like there would be certain people that you would want them to be less explicitly focusing on their emotions.

DR. LANE: Absolutely! So one of the things; I mean what immediately comes to mind is recent work on rumination. I mean the field is really expanding greatly and in the past year there's been a fMRI study of rumination. And what it shows is that rumination is associated with first decreased frontal lobe activity, and secondly increased and prolonged amygdala activity. So I mean another issue that comes up in response to that is that women are consistently, you know, they score higher on the LEAS, and I think in general are more aware. But we also know that women have twice the rate of major depression and various anxiety disorders than men do. And when is it an advantage? When does it become a liability? I think issues such as rumination become relevant and there must be some kind of dysregulation that occurs if you're too focused on your feelings.

DR. GRANDEY: I was just thinking about the ideas of implicit processes and you talked about being aware of your emotions based on your facial expressions. I wondered if you could just say your opinion or your view on the facial feedback hypothesis. Do you know what I'm referring?

DR. LANE: Yes. Oh, sure. My opinion about the facial feedback hypothesis is that there's a real effect there, but that it; and I would say that Damasio's work has really been influential in getting me to think that there really is something important about bodily feedback. But I think that it's only part of the story, and I think it's probably a relatively minor effect, although I do know; I'm familiar with the studies that say that if you put your face in a particular configuration without knowing that it's emotional that it will be associated with certain subjective states and you will get EEG changes that are consistent with the affective state that you're trying to create. I think there's something to it, but I think it's just one part of a larger picture.

DR. GRANDEY: But the bodily, and I might've missed some stuff too, but being aware of your bodily changes was one of the first stages of your awareness model.

DR. LANE: That's true.

DISCUSSION

DR. GRANDEY: And it's an implicit process you're not aware of – correct?

DR. LANE: Yes.

PARTICIPANT: So that's more of the facial feedback.

DR. LANE: Yes, that is.

PARTICIPANT: So your model goes beyond…?

DR. LANE: I think there's something more than just; I mean Damasio would say a feeling is your conscious awareness of your bodily state, and I think that there may be something more to it than that, but I think that that's part of it certainly. And I would; well, who am I to disagree with Damasio?

DR. GRANDEY: I have my own problems with it. Maybe we could talk later. I don't need to carry on.

DR. LANE: Okay. But I think what you're implying is that the facial feedback hypothesis is entirely consistent with my model.

DR. JOHNSON: So how do you place in your levels emotions with complex cognitive structures such as guilt, remorse, dread, things like that?

DISCUSSION

DR. LANE: Good, no make that an excellent, question. We do not differentiate between those kinds of emotions that you just described and simpler emotions like happy, sad, anger, fear, love, hate. And that's an example of an aspect of the scoring system that we haven't addressed, and it would be nice to go back and see whether if we somehow took those more complex emotions into account, whether that would add to our predictive power of other things. One of the things that I want to be sure to say is that Kim Barchard, who's been mentioned several times, who did her dissertation on emotional intelligence, has developed a computerized scoring system for Levels of Emotional Awareness Scale, and she says that in their most advanced algorithm that the correlation with hand scoring is .85. I think when we have a computerized data base and can interrogate the data efficiently we can address questions like that.

DISCUSSION

DR. MATTHEWS: I have a question concerning automaticity, because I think this is a construct that is easily misused and especially learned automaticity. And I think the stick shift example which you gave is constructive here, because traffic psychologists – I dabble in traffic psychology – have actually debated this issue. And it seems that stick shifting, even among highly experienced drivers, is not fully automatic. It requires some cognitive resources. And part of the reason for that is that that behavior has to be sensitive to the motivational context, because in general you shift down when you're slowing down. But occasionally you want to pass a vehicle, or you want to burst pass, so you shift down to do that. It's serving a different goal. And perhaps the same thing applies to emotions. For example, if you're expressing anger there are different tactics which you may be using depending on the context. You might want to express anger to show empathy with a friend who's frustrated about something. Or you might want to express anger more in order to express aggression to someone. So it's important that response doesn't get fully automated, because it has to be sensitive to these contextual factors and to personal motivations.

DISCUSSION

DR. LANE: And I think what you're talking about is the difference between someone who's highly emotionally intelligent and someone who isn't. I think somebody who's at a lower level of emotional intelligence or emotional awareness isn't going to be able to shift gears and recognize that the situation is different and change their automatic behavior. And I think the reason why mental health clinics are filled with lots of people is that they have maladaptive behavior patterns that they don't know what to do with. And I think what you're describing is somebody who really can go in and out of, and can monitor their automatic behavior and switch out of it if need be.

DR. ROBERTS: Thank you very much, Richard. As you can see we've finished the innings and are now off to dinner.

Editors Comments

This is the end of Volume II. Proceedings are continued in Volume III.