

Conveying Affect: Vocal vs Nonverbal Cues

Joshua Y. Zheng

Oregon State University

Author Note

This report was written in Winter Term 2016 for PSY 460, Advanced Social Research Methods, instructed by Dr. Frank Bernieri. This report was nominated by Dr. Frank Bernieri for the 2016 OSU Culture of Writing Award in Psychology.

April 21st, 2016.

### Abstract

This study was interested in determining whether vocal communications or facial expressions and body language are more effective in communicating positive, neutral or negative affects when used in isolation and without deception. A sample of six male and six female (n = 12) students in Dr. Bernieri's Winter term 2016 PSY 460 class were taken and randomly assigned to an audio (vocal) or visual (facial and body language) condition. A video of 189 communications was split into either only the audio or only the video, and played back to the audio and visual condition respectively. Participants were asked to judge what affect was being communicated for all 189 communications, and afterwards were given three personality tests measuring social skills and interpersonal sensitivity. It was found that the participants in the audio condition were able to decode messages more accurately than the visual condition, and that there was no significant difference in decoding accuracy between men and women across both conditions.

*Keywords:* Affect, channel, decoding, non-verbal cues, visual primacy

### Conveying Affect: Vocal vs Nonverbal Cues

For the most part we all need to communicate with our fellow human beings on a regular basis to get things done, thus it makes sense for us to know what we should be looking for when trying to understand how someone is feeling. Is it the words they say or the tone of their voice that conveys feeling? Perhaps their facial expression, body language or hand gestures carry more weight than those verbal cues. Furthermore, we could ask whether some feelings are better communicated vocally instead of non-verbally or vice versa. These are all questions that need to be answered for us to better understand what makes for good communication assuming most people are not out to lie to us or deceive us.

Many studies have been conducted on the subject of whether speech (tone of voice, the words we say) is more effective than visual channels (our body language, facial expressions, gestures) at communicating emotions. A theory that has been created based on research on this subject is the theory of visual primacy. Visual primacy is the idea that we tend attach more weight to visual cues (what we see) than audio cues across a wide range of perception and memory tasks (Posner, Nissen & Klein 1976). The theory of visual primacy gave rise to Mehrabian and Ferris' "7-38-55" rule of communication (1967) – a very commonly cited rule in popular media regarding communication. The rule purports that if you were to weight how well words, tone of voice and facial cues communicate how someone is feeling, you would find that tone of voice accounts for 38%, facial cues are responsible for 55%, and words mean almost nothing at a measly 7%. What this research implies is that when trying to tell how someone is feeling we should primarily rely on body language and their tone of voice – but given the choice we should prioritize body language over tone of voice. Mehrabian's work is supported by other investigators who found that the visual channel is dominant at a degree that is significantly larger

than the vocal channel. In comparison, the verbal channel was again found to be the least effective at conveying emotional meaning (Burns & Beier 1973).

On the other hand, research has also shown that people who pay attention to non-verbal cues in addition to vocal cues are able to identify emotions much more accurately than those who only use a verbal component (Archer and Akert 1976). Work done by Archer and Akert had participants either watch a video that contained multiple scenes of people having conversations (where facial expressions, movement, varying tone of voice were all discernible) or read a text transcript of that video before answering interpretative questions. For example, participants would either watch or read a conversation between two men who just finished a basketball game and then would be asked to interpret which of the two men won the match. The difference in accuracy between the text only and video group was so profound that those who only used a verbal component (words only, no tone of voice or other non-verbal cues) performed significantly worse than if they had blindly guessed the emotions being expressed. Based on this research we can conclude that non-verbal cues are indispensable in our ability to interpret how someone is feeling.

Going a step further, even in cases where we are specifically told to only focus on tone of voice research has shown that we are invariably influenced by our perception of the speaker's whole body (Stock, Righart and Gelder 2007). The fact that we are so drawn to visual cues when trying to determine affect even when specifically told not to lends credit to the theory of visual primacy. The visual channel of communication is dominant even in day to day conversations – such as telling how someone is being hurtful or humorous (Haase & Tepper 1978). Jokes and sarcasm rely on discrepancies between communicative channels to be expressed properly, so in most cases require three primary components that need to be weighted in conjunction by the

decoder: tone of voice, words and facial expression (Bugental, Kaswan & Love, 1970). Various combinations of positive or negative affect expressed by the three different channels could represent completely different affective meanings, but research by Bugental and company indicates that the strength of the visual channel (facial expressions) in decoding conflicting messages is almost twice that of either the vocal or verbal channel. A similar effect was found in work done by DePaulo and Rosenthal (1978), where it was found that participants tended to pay more attention to the visual modality when trying to decode affect given contradicting visual and vocal cues (e.g. positivity in one cue, negativity in the other). While these results can't necessarily be generalized to sincere communications as the experimental conditions involved deception, these experiments demonstrate that the visual channel tends to be dominant and effective when trying to decode affect overall.

A problem to consider is that our tendency to attach more weight to visual cues does not necessarily tell us whether or not doing so is actually the most effective way to tell how someone is feeling. For example, it has been shown that the face is better at communicating affect, whereas the voice is better at communicating dominance or submission. This means that if you were trying to figure out how someone was feeling (e.g. positivity, negativity, like, dislike etc.) using only verbal or non-verbal channels you would be better off looking at their face compared to listening to their tone of voice (Ekman 1965; Zuckerman et al. 1975). This effect may account for why misinterpretations can result when communication is restricted to a less or single channels like that of email, text messaging or phone calls.

Some research has even found that in certain cases the influence of the vocal channel is dependent on what is being conveyed in other non-verbal channels. A study done by Argyle et al. (1971) discovered that a positive tone of voice has no effect when it is paired with hostility

conveyed through non-verbal cues. The results found by Argyle and company contradict the work done by Mehrabian and Ferris (1967) who instead concluded that there was no interaction between facial and vocal components. This is important to note because evidently the two channels *can* interact and affect each other which makes Mehrabian's 7-38-55 model questionable and/or difficult to generalize to all communications.

More recent research has tackled this notion of channel superiority with a more moderate position of multisensory perception of affect (MPA). MPA asserts that no channel is dominant over the others in all cases, but that different contexts/situations will lead to advantages for focusing on a particular channel (Collignon et al. 2008). MPA predicts that there are "flexible, situation-dependent rules that allow information to be combined with maximal efficacy", and that this framework is what dictates which channel is prioritized in cases where the other channels are sending unclear or ambiguous messages. This means that rather than focusing on the relative strengths of each channel it is more important to view them in context of the situation. What this means for the present study is that the experimental design should be the biggest predictor of which channel of communication is most powerful. We should then, for example, expect to see the visual channel be more effective if deception were to be involved or see the vocal channel perform better if the visual information presented is unclear or ambiguous.

Beyond the various channels of communication, we can also look at individual factors such as gender which could potentially influence how well we are able to interpret how someone else is feeling, but also how well we are able to convey how we are feeling to other people as well (i.e. encoding vs decoding). For example, maybe one gender is more sensitive to non-verbal cues and can interpret them better than the other gender, or maybe one gender is more expressive and can communicate in such a way that anyone can easily tell how they are feeling. The former

question is answered by research done by Hall (1978) who found that women tend to be better than men at interpreting emotions. The strength of the effect found translated to the “upper 50% of the female distribution exceed[ing] about 65% of the male distribution” which is an advantage that cannot be ignored when conducting any research on interpretation of emotion. This effect is corroborated by research conducted by Rosenthal and DePaulo (1979) as a similar advantage for women was found.

In terms of expressivity, research by Noller (1980, 2001) provided further evidence that women are better at decoding than men, but also better at *encoding*. In Noller’s study, married couples were recruited and each assigned one of two possible sets of 27 cards: an encoding set or a decoding set. Both sets presented a situation on each card for the spouse to act out (e.g. a cold winter’s night), but the encoding set included an intention (positive, negative, neutral) and a specific sentence to say. The decoding set also contained the situation but instead included three interpretations of what the spouse was trying to convey. It is important to note that by keeping the sentences that the spouses had to read each other constant across all communications, Noller was able to control for the effect of verbal causes such that any variations in affect had to be attributable to non-verbal causes (standard content methodology). Each trial consisted of the encoder reading off the written statement in a manner consistent with the written intention and the decoder marking what they thought the intention was. Analysis of the results indicated that for the most part women were superior at both encoding messages, particularly for positive messages, and decoding them as well. Noller effectively found that it’s easier to tell what women are feeling based on their non-verbal cues compared to men, and that women in general are better at telling how other people are feeling.

Several sources have found a distinct, significant advantage that the visual channel has over the vocal and verbal channel. Mehrabian and Ferris (1967) found that the visual channel had the most weight, and their results were supported by the work done by Burns & Beier (1973) who found a similar, significant advantage for the visual channel over the vocal channel as did various other investigators (Bugental, Kaswan & Love, 1970; DePaulo & Rosenthal 1978). Analysis of the literature available also indicates that women have an advantage over men in both encoding and decoding affect (Hall 1978; Noller 1980, 2001). Therefore, given the above, this study hypothesizes the following:

1. The visual channel will convey positive, neutral and negative affect more accurately than the vocal channel.
2. Women in this study should decode affect better than men.

## **Methods**

### **Participants**

The participants were comprised of six male and six female students in the Winter 2016 section of PSY 460 taught by Dr. Frank Bernieri. All students in the section were upperclassmen Psychology majors at Oregon State University. All students in the section opted to participate and were each assigned a number. Six students were randomly assigned to the audio condition, and the rest were assigned to the visual condition (both of which will be described below).

### **Materials and Design**

Materials consisted of a video projector and two prepared videos, a packet with 189 questions and a packet with three personality tests. In combination, the two prepared videos and packet of 189 questions comprise the Emotional Judgment Task (EJT). The Emotional Judgment



Task is a test that was designed by Dr. Frank Bernieri at Oregon State University based off of a standard content task developed by Patricia Noller from the University of Queensland (Noller 2001). The general idea behind standard content methodology is that first you create scenarios that an individual could encounter in day to day life, such as sitting in a classroom with a friend. Then, you create an ambiguous line of dialogue for each scenario (“What do you think of this teacher?”). “Ambiguous” in this case meant that the line needed to have different meanings depending on the nonverbal cues it was said with (i.e. the line can be used to convey positive, negative or neutral affect). Done properly, you can then assume that any changes in the affect conveyed by the message must be caused by nonverbal channels and not the verbal/words (i.e. the verbal channel is controlled). Our experiment took the procedure one step further and separated the vocal channel from the visual channel using the aforementioned Emotional Judgment Task.

To design the EJT, nine scenarios (e.g. Peanut Butter, Class, Movie, Dishwasher etc.) were created and each scenario was assigned an ambiguous line of dialogue. A group of students from Oregon State University were recruited and split into groups of seven. Each student cycled through the scenarios and performed the line associated with it three times in either a positive, neutral or negative way. These performances were recorded in two second video clips. In total, there were three affects and nine scenarios making for 27 possible combinations. The 27 combinations were repeated seven times by different students thus resulting in 189 two second video clips. The clips were edited into two videos: one in which only the voice of the actors could be heard (verbal tape), and one in which only the body language and facial expressions of the actors could be seen (non-verbal tape). In the audio tape participants would see a still frame with the name of the scene and a number assigned to it (Example: 1. Class or 15. Peanut Butter)

before listening to a two-second audio clip. The non-verbal tape was identical, except instead of a two-second audio clip the participants saw a silent video of an actor.

A packet titled 'Content Standard Video' was created and had the instructions "Please circle the letter of the answer or intention of the speaker that you feel they are trying to portray." The 189 questions matched the order of the scenes presented in the two edited videos and example questions and responses can be found in the appendix. The key word that distinguished the emotion in the positive and negative response was bolded. This packet was the same for both conditions. How the test was administered will be described in the 'Procedure' section below.

The personality questionnaires were comprised of the Riggio Social Skills Inventory, the Toronto Alexithymia Scale and the Situational Test of Emotion Management (STEM). The Riggio SSI was designed to assess seven different social skills such as emotional expressivity, emotional sensitivity, social control and social sensitivity (Riggio 1986). The version of the Riggio given only had 90 questions and examples of questions on the Riggio would be: "Few people are as sensitive and understanding as I am" or "It is difficult for others to know when I am sad or depressed." The responses for each question were on a 1-5 rating scale with 1 representing a personal judgment of "Not at all like me" to 5 which represented "Exactly like me." The SSI has been significantly correlated with traditional scales of personality and communication skills such as the 16 PF. In fact, measures of emotional sensitivity on the SSI correlate strongly with PF dimensions of outgoing and sensitive individuals which were linked to ability to tell how an individual is feeling which predicts that those who score high on the SSI should perform better on the EJT.

The Toronto Alexithymia Scale (TAS) is a 20 question test designed to diagnose alexithymia: a condition where an individual has difficulty or is unable to experience, express or

describe emotions (Bagby, Parker and Taylor 1994). It is assessed with three factors: difficulty identifying feelings, difficulty describing feelings, and externally oriented thinking. Examples of questions on the TAS include “I am often confused about what emotion I am feeling” and “I am able to describe my feeling easily” with responses on a 1-5 rating scale with 1 representing a personal judgment of “strongly disagree” and 5 representing “strongly agree.” In a follow up paper, Bagby, Parker & Taylor (1994) described the convergent and discriminant validity of the TAS-20 and mentioned that the test was very negatively related to various scales measuring openness to experience, feelings and fantasies – all of which are consistent with the construct of alexithymia. The most striking test was that the investigators had those who scored very high on the TAS-20 take a test to see how well they could recognize facial expressions of emotion, and found that they scored less than average. In essence, the TAS was administered because one of the consequences of alexithymia is also the inability to tell how someone else is feeling which could potentially result in an afflicted participant getting very low scores on the Emotional Judgment Task (particularly in the visual condition).

The Situational Test of Emotional Management (STEM) is a 44 question test that assesses how well someone responds to emotionally stressful situations (MacCann and Roberts 2008). It was designed as a potential measure of emotional intelligence. An example of a question and the responses associated with it would be:

Lee’s workmate fails to deliver an important piece of information on time, causing Lee to fall behind schedule also. *What action would be most effective for Lee?*

- (a) Work harder to compensate.
- (b) Get angry with eh workmate
- (c) Explain the urgency of the situation to the workmate.

(d) Never rely on that workmate again.

This test was administered because high scores may be correlated with high scores on the EJT and the Riggio SSI.

### **Procedure**

The class was informed that the purpose of this study was to compare whether it was easier to tell how someone was feeling based on the tone of their voice or their body language and facial expression. The visual and audio groups were then directed into two different rooms and each given the packet of 189 questions. Both groups were instructed to be seated, ensure they were able to see the screen in the visual condition or hear the video in the audio condition, and read the instructions on the Content Standard Video packet given. Participants were instructed to view/listen to each video segment and circle the answer they believed best represented how the actor was feeling. After the video was complete, the packet was collected and the personality questionnaires were handed out to each participant<sup>1</sup>. Upon completion of the personality questionnaires, participants were thanked for their participation and allowed to leave. Participation in this experiment took about fifty minutes.

### **Results**

It was expected that the participants in the audio condition would on average correctly judge affect better than the visual condition, and that female participants across all conditions would be better judges than males. A comparison of the descriptive statistics for the audio and visual condition can be found in Table 1.

---

<sup>1</sup> In the audio condition there was a technical difficulty with the laptop used to show the video. Due to time constraints the participants in the audio condition had to take the personality questionnaire before they took the EJT. Implications of this will be assessed in the Discussion section.

The audio condition and visual condition each had six participants ( $n = 6$  for each condition) evenly split between male and females (three males and three females per condition). The grand mean for the audio condition was 112 and the grand mean for the visual condition was 104 which means that those in the audio condition on average scored eight points high than those in the visual condition, and furthermore, this difference is significant:  $t(9) = 4.07, p < 0.05$ .

Males in the audio condition had a mean of 113 ( $\bar{x} = 113$ ) with a standard deviation of 1.53 (SD = 1.53), the difference being negligible in comparison to females in the audio condition who had a mean of 111 ( $\bar{x} = 111$ ) with a standard deviation of 4.73 (SD = 4.73). Males in the visual condition had a mean of 102 ( $\bar{x} = 102$ ) with a standard deviation of 4.16 (SD = 4.16) whereas females in the visual condition had a mean of 106 ( $\bar{x} = 106$ ) with a standard deviation of 2.08 (SD = 2.08). The females performed better than the males at a degree that was a little over two times greater than what would be expected, but this difference was not significant:  $t(7) = -2.11, p = 0.073$ .

A two-way ANOVA was performed and the results can be found in Table 2. As mentioned earlier, a significant effect was found for channel of communication,  $F(1,8) = 17.27, p < 0.05$  but no significant interaction effect was found,  $F(1,8) = 2.08, p = 0.187$  nor was there a significant effect found for sex  $F(1,8) = 0.56, p = 0.58$ . Overall, this study found that there was a significant advantage for the audio condition in terms of their ability to judge affect and this study did not find any difference between the two sexes in terms of decoding ability.

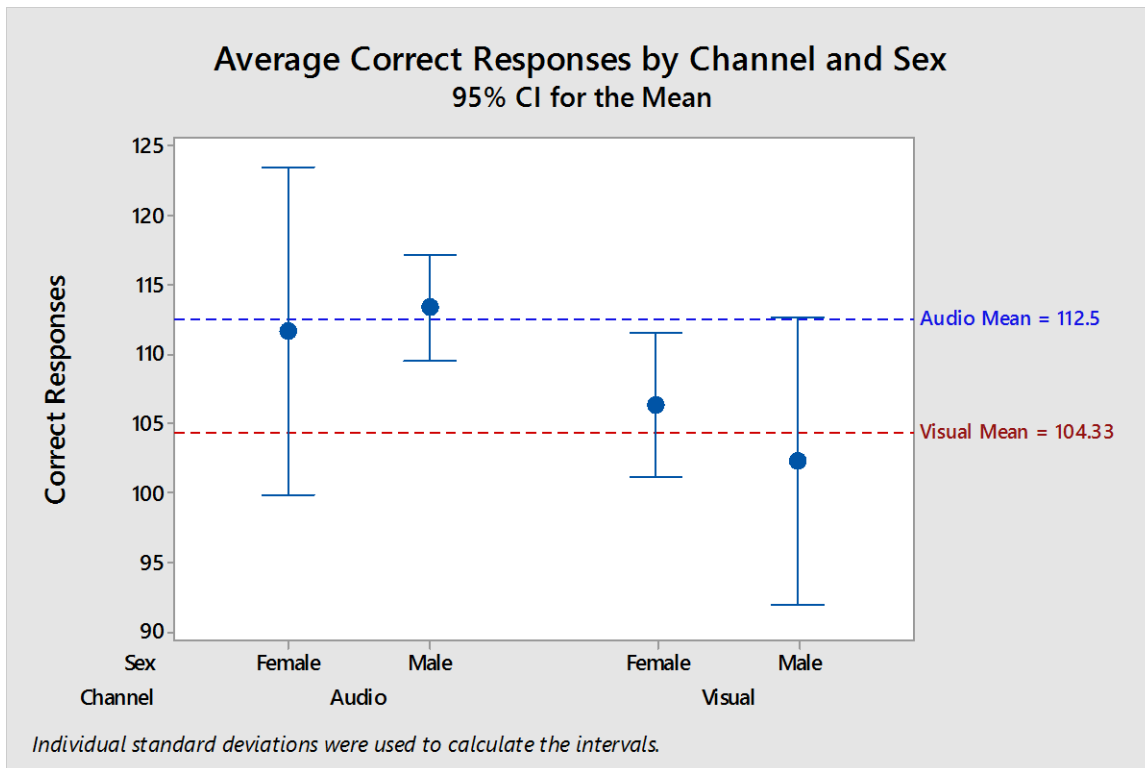


Figure 1. Average number of correct responses on EJT for each condition and sex. The two reference lines refer to the grand means for the audio and visual conditions, and the 95% confidence intervals are for the mean per sex group in each condition is displayed.

### Discussion

Based on the literature presented earlier (Burns & Beier 1973; Bugental, Kaswan & Love 1970; DePaulo & Rosenthal 1978; Mehrabian & Ferris 1967; Posner, Nissen & Klein 1976) this study expected to find that the participants in the visual condition would be able to judge affect better than those in the audio condition and that women would score better than men overall on the EJT. Analysis of the results did not support either hypothesis, as a significant main effect for channel was found in favor of the audio condition and no significant difference in performance was found between males and females (refer to Figure 1 for a visualization).

One potential explanation for the results is that the communicative messages we filmed were not the result of natural emotional states. There is a potential ethical issue in inducing certain strong emotional states (anger, sadness, any negative affective state etc.) which means we had to film communications where the intensity of the affect being felt by the actor was relatively low. What this means is that the affect conveyed in each case might not have been 'emotional' enough such that the participants who had to decode them later were easily able to discriminate between them (Scherer 1986). There may have been, for example, significant overlap between the positive and neutral or neutral and negative messages such that communications we presented could have been reasonably interpreted as being potentially multiple affects versus one distinct one. This is not to say that the communications used in this study were useless due to their inhibited nature and cannot be used to make any inferences about reality. The communications are just limited in that they only inform us about situations where someone is trying to convey a certain affect whereas they might be feeling something completely different (acting, being polite, lying etc.).

Another factor that could help explain the results we obtained is that the environment created for the participants does not reflect the vast majority of day to day communications. All the encoded messages were pre-recorded and played back at the participants which could have resulted in a loss of critical visual or audible information that could have been used to more accurately decode affect (Bachorowski 1999). For example, perhaps the fact that we used a video reduced cues like depth perception or clarity of facial features in a way that gave the audio condition an advantage. Most of the time, when we communicate with people, we are interacting with them face to face and thus the messages we receive are not distorted by technology. On the other hand, it isn't too uncommon for people to make a few phone calls a day or use video

conferences in the corporate environment so these results could tentatively be used to inform us of those mediums of communication. In summary: given the design of our experiment we should focus in part on generalizing the effects we found to those cases where technology is aiding communication and *not* most cases of day to day, face to face communication.

Another possibility is that all of the participants in this study, including the ones who were used to create the communicative messages, were psychology majors. Psychology majors, in addition to being pre-dominantly from a certain age demographic, tend to have participated in a lot of previous research which can cause them to be fearful of deception and thus present themselves in a way that is not natural or characteristic of the larger population (Scherer and Ekman 1982b). This could have biased the messages that were recorded such that they were even less representative of natural communication than they already were.

After a review of the literature, the most compelling argument to explain why the results found by the present study differed so drastically from the hypothesized effects has to do with the problem that many of the studies this investigation used to formulate its hypothesis (Bugental, Kaswan & Love 1970; Mehrabian & Ferris 1967; Posner, Nissen & Klein 1976; DePaulo & Rosenthal 1978) were not meant to be generalized to the experimental design used in this study. The hypothesis of this study was heavily influenced by research on the visual primacy effect which was the idea that visual cues tend to be weighted more than audio cues (Posner, Nissen & Klein 1976). However it would seem that the tendency to weight visual cues mostly occurs when the verbal message it is paired with is weak, unclear or discrepant (Bugental et al. 1976, Zuckerman et al. 1982). Looking at studies where the verbal component of speech is minimized or removed (DePaulo 1977, Mehrabian & Ferris 1967) through methods such as content filtering, the result is that participants pay more attention to facial expressions over



speech cues. In essence, when participants are forced to choose between facial expression and garbled audio nonsense, they choose to use facial expression to figure out what the affect is. This makes sense because content filtered speech is unlike anything most people have heard in their lifetimes and only a vague semblance of what we consider to be normal human speech.

Unaltered facial expressions on the other hand are familiar to us and we tend to all have our own ways to interpret facial expressions whereas few individuals, if any, have ever had to come up with a strategy to understand pure tone of voice devoid of meaning or words. That being said, the present study only attempted to control the verbal cues such that any changes in affect could only be attributed to vocal cues. The audio condition contained a coherent, understandable communication with words in stark contrast to the work done by Mehrabian and DePaulo. Therefore, the fact that the verbal cues were intelligible and clear may have produced an interaction with the vocal cues that can explain why the participants in the audio condition scored better than the visual condition.

It is useful to view this in the context of the theory of leakage hierarchy (Ekman et al. 1980) and its relation to DePaulo's 1977 study. Leakage hierarchy implies that the vocal channel is the leakiest channel and thus the one that is most likely to give away information about what the encoder is trying to hide, and DePaulo's 1977 study demonstrated that in cases where individuals sense that they are being deceived they have a tendency to focus on audio cues (verbal and vocal). Since the participants in the present study were aware that all messages they were asked to decode were generated through actors, acting being a form of deception, it makes sense that those participants in the audio condition had an advantage over those in the visual condition since they could focus on the channel they would have attended to anyway. What this means is that the most effective decoding strategy given this study's experimental design was to

focus on voice: the leakiest channel. The advantage for doing so may have been large enough to account for why the audio condition outperformed the visual condition.

Before discussing implications of this result, it's important to re-iterate that the results of this study should only be generalized to specific types of day to day communication – those involving a case where the encoder is attempting to convey an affect that they may not actually be feeling. The results of this study can broaden our understanding of situations that require acting/deception or taking on a highly scrutinized role (e.g. therapists, teachers etc.). It also lets us learn more about specific interactions with strangers, colleagues or friends and family, for example: when one is attempting to convey a positive attitude to a professor out of respect, when in fact they may actually be feeling neutral or negative about said professor.

The results of this study suggest that vocal cues are more easily decoded and in fact preferred over visual ones. What this means is that people may have a strong tendency to focus on vocal cues and verbal content when speaking to individuals in occupations such as counseling or marketing where the role by nature involves needing to suppress certain affects in favor of others (e.g. positivity, affection, good-will). They may also do this in general as a rule of thumb simply because the vast majority of daily interactions for socialized individuals must be conveyed as positive regardless of what one is actually feeling. Expressing negative or neutral affect tends to be frowned upon. As such, it may be beneficial to train people to be aware of how much information can be leaked vocally and rely primarily on non-verbal communication to convey affect instead where possible.

That being said, the results raise some questions about the experimental design. Some literature found that no one channel of communication is dominant over the others (Ekman et al. 1980; Krauss et al. 1981), and the researchers that did find dominant results only did so because

their experimental designs were artificial (DePaulo 1977, Mehrabian & Ferris 1967). If the experimental design truly was the cause of the results we obtained, we could verify this with another experiment. A direction for future study could be to take all of the speech recorded in the present study, content filter them such that only tone of voice is left, and then re-administer the EJT such that the visual condition is left the same but the audio condition no longer has any verbal content at all. As discussed earlier, work done by Mehrabian and DePaulo (DePaulo 1977, Mehrabian & Ferris 1967) used strange communicative messages where the audio used was completely devoid of linguistic meaning whereas the present study used audio that still had linguistic meaning. It was hypothesized that perhaps the reason the present study's results differed from previous literature was specifically because of this key difference. Content filtering the communications in this study would bring the experimental design in line with the designs used by DePaulo and Mehrabian, and therefore, doing so could potentially reverse the results found in this study and cause the visual condition to decode more accurately than the audio condition as originally hypothesized. Another question that could be posed going forward would be how well vocal cues communicate more discrete emotions. While the present study operated with the relatively broad terms of "positive, neutral and negative" affect, it may be of interest to focus on other emotions such as affection, disinterest or hatred and see if vocal cues are still more effective than visual cues at communicating such states.

At this time it would be good to touch on some further limitations of this study. One of them was a technical failure when the EJT was being administered in the audio/upstairs condition. Originally, the audio condition was supposed to be taking the EJT at the same time as the visual condition downstairs, but a fault in the laptop that was being used to play the communicative messages delayed the task. The audio condition was instead instructed to

complete the personality questionnaires and return in two days' time to take the task itself. The major problem is that the personality questionnaires were self-report and asked questions related to how socially/emotionally sensitive the participant believed themselves to be. This had a high potential for creating an issue with demand characteristics. For example, if a participant were to have self-reported that they were very socially aware and sensitive, they would have a need to perform well on the EJT in an effort to live up to that expected standard which could have influenced how much they deliberated each item on the EJT and how much focus or effort they put into taking the test. A similar bias could be argued for those who rated themselves poorly on the questionnaires.

Another major limitation was that this study was conducted with a PSY 460 class where all of the participants were made very aware of the nature of the study and some of the research informing the question being posed by the study which may have created bias since participants likely already formed their own opinion of what the results of the study should have been. Thus the samples used in this study were convenience, and the "random sampling" conducted by Dr. Bernieri was hardly random at all. Furthermore, of the 20+ samples that were collected only 12 were ultimately used in the statistical analyses and they were handpicked by Dr. Bernieri using unknown selection criteria. It's entirely possible that he chose specific samples to get the result the present study obtained although this is unlikely.

Communication is a necessary, complex process. Understanding how someone feels is more than just focusing on their words, face or voice, and there does not seem to be a clear champion when it comes to which of these many modes of communication is the best at either getting across what we feel or understanding how other people feel. There is no 'secret ratio that applies to all situations like the one developed by Mehrabian (Mehrabian & Ferris 1967), as

communication is far too complex to be reduced to a catch-all linear relationship. However, this study has helped to broaden our understanding in that it found that we are very likely better off listening to what someone is saying and the tone of voice with which they say it compared to just looking at their face or body language especially when we suspect deception. In fact, we find deception everywhere in that most of the people we encounter are all out to convey only positive affect whereas the reality may very well be that most people are feeling neutral or even negative most of the time. The results of this study indicate that you are likely to judge how someone is feeling incorrectly if you only attend to the smile on their face and not their voice. Overall though, the context of all communication is key and depending on the situations we find ourselves in it might just be better to look rather than listen or vice versa.

### References

- Archer, D., & Akert, R. (1976). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology*, 35(6), 443-449.
- Argyle, M., Alkema, F., & Gilmour, R. (1971). The communication of friendly and hostile attitudes by verbal and non-verbal signals. *European Journal of Social Psychology Eur. J. Soc. Psychol.*, 1(3), 385-402.  
<http://dx.doi.org.ezproxy.proxy.library.oregonstate.edu/10.1002/ejsp.2420010307>
- Bagby, R., Parker, J. D., & Taylor, G. J. (1994). The Twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23-32.
- Bagby, R., Taylor, G. J., & Parker, J. D. (1994). The twenty-item Toronto Alexithymia scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, 38(1), 33-40. Retrieved February 28, 2016.
- Bugental, D. E., Kaswan, J. W., & Love, L. R. (1970). Perception of contradictory meanings conveyed by verbal and nonverbal channels. *Journal of Personality and Social Psychology*, 16(4), 647-655. Retrieved March 9, 2016.
- Burns, K. L., & Beier, E. G. (1973). Significance of Vocal and Visual Channels In the Decoding of Emotional Meaning. *J Communication Journal of Communication*, 23(1), 118-130. Retrieved March 9, 2016.
- Bachorowski, J. (1999). Vocal Expression and Perception of Emotion. *Current Directions in Psychological Science Current Directions in Psychol Sci*, 8(2), 53-57.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F.

(2008). Audio-visual integration of emotion expression. *Brain Research, 1242*, 126-135.

Retrieved February 28, 2016.

Depaulo, B. M., & Rosenthal, R. (1978). Decoding discrepant nonverbal cues. *Journal of*

*Personality and Social Psychology, 36*(3), 313-323. Retrieved March 9, 2016.

Ekman, P. (1965). Differential communication of affect by head and body cues. *Journal of*

*Personality and Social Psychology, 2*(5), 726-735.

Ekman, P., Friesen, W., O'sullivan, M., & Scherer, K. (1980). Relative importance of face, body,

and speech in judgments of personality and affect. *Journal of Personality and Social*

*Psychology, 38*(2), 270-277.

Haase, R., & Tepper, D. (1972). Nonverbal components of empathic communication. *Journal of*

*Counseling Psychology, 19*(5), 417-424.

Hall, J. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*(4), 845-

857.

Krauss, R. M., Apple, W., Morency, N., Wenzel, C., & Winton, W. (1981). Verbal, Vocal, and

Visible Factors in Judgments of Another's Affect. *Journal of Personality and Social*

*Psychology, 40*(2), 312-320.

MacCann, C., & Roberts, R. D. (2008). New Paradigms for Assessing Emotional intelligence:

Theory and Data. *Emotion, 8*(4), 541-551.

Mehrabian, A., & Ferris, S. (1967). Inference Of Attitudes From Nonverbal Communication In

Two Channels. *Journal of Consulting Psychology, 31*(3), 248-252.

- Noller, P. (1980). Misunderstandings in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology, 39*(6), 1135-1148.  
Retrieved January 30, 2016.
- Noller, P. (2001). Using Standard Content Methodology to Assess Nonverbal Sensitivity in Dyads. In F. J. Bernieri & J. A. Hall (Eds.), *Interpersonal Sensitivity: Theory Measurement and Applications* (pp. 243-264). Lawrence Erlbaum Associates.
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*(2), 157-171.  
Retrieved March 2, 2016.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51*(3), 649-660.
- Rosenthal, R., & Depaulo, B. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology, 37*(2), 273-285.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*(2), 143-165. Retrieved February 22, 2016.
- Scherer, K. R., & Ekman, P. (1982b). Methodological issues in studying nonverbal behavior. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 1 -44). Cambridge, England: Cambridge University Press
- Stock, J., Righart, R., & Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7*(3), 487-494.  
<http://dx.doi.org.ezproxy.proxy.library.oregonstate.edu/10.1037/1528-3542.7.3.487>



Zuckerman, M., Lipets, M., Koivumaki, J., & Rosenthal, R. (1975). Encoding and decoding nonverbal cues of emotion. *Journal of Personality and Social Psychology*, 32(6), 1068-1076.

Zuckerman, M., Amidon, M., Bishop, S., & Pomerantz, S. (1982). Face and tone of voice in the communication of deception. *Journal of Personality and Social Psychology*, 43(2), 347-357.

**Table 1.** *Descriptive Statistics for Sex, Condition (Audio/Visual) and Decoded Messages*

Sex	Channel		Grand Mean
	Audio	Visual	
Male	$\bar{x} = 113.3$ n = 3 SD = 1.53	$\bar{x} = 102.3$ n = 3 SD = 4.16	107.8
Female	$\bar{x} = 111.6$ n = 3 SD = 4.73	$\bar{x} = 106.3$ n = 3 SD = 2.08	109
Grand Mean	112.5	104.3	

**Table 2.** *Two-Way Analysis of Variance (ANOVA) Between Communicative Channel and Sex*

Source	SS	df	MS	F	p
Between Groups	228.24	3	76.08	6.56	0.015
Sex	4.08	1	4.08	0.56	0.58
Channel	200.08	1	200.08	17.27	0.003
Sex*Channel	24.08	1	24.08	2.08	0.187
Error	92.66	8	11.58		
Total	320.91	11			

## Appendix

### *Data Used For Analysis*

There were 189 items. 63 positive messages. 63 negative messages. 63 Neutral messages.

0=Male  
1=Female

1=Silent Visual condition (downstairs)  
2=Audio only condition (upstairs)

$N = 12$   
 $n = 3$  within each condition defined by sex and channel.

<b>Obs</b>	<b>sex</b>	<b>channel</b>	<b>Total Correct</b>	<b>Correct Positive</b>	<b>Correct Negative</b>	<b>Correct Neutral</b>
<b>1</b>	0	1	101	23	49	29
<b>2</b>	0	1	107	32	49	26
<b>3</b>	0	1	99	37	34	28
<b>4</b>	1	1	108	36	44	28
<b>5</b>	1	1	107	28	48	31
<b>6</b>	1	1	104	35	40	29
<b>7</b>	0	2	115	34	48	33
<b>8</b>	0	2	113	29	46	38
<b>9</b>	0	2	112	33	49	30
<b>10</b>	1	2	110	40	40	30
<b>11</b>	1	2	108	37	43	28
<b>12</b>	1	2	117	37	49	31

*The following is an example of the questions and responses that participants used to complete the task.*

- |  |  |   |
|--|--|---|
| <p>1. Peanut butter</p> <p>You go to your refrigerator and notice that there is no more peanut butter; your roommate is cooking dinner nearby.</p>   | <p>"Hey, there's no more peanut butter."</p>               | <p>a. Your roommate is going to the store this afternoon and you are relieved/happy you got a chance to add PB to the shopping list.</p> <p>b. You're mad because your roommate always eats your peanut butter without leaving you any.</p> <p>c. Your roommate likes peanut butter too, and you just want them to know there's none left in case they're planning on having some</p> |
| <p>2. Surprise Party</p> <p>You come home and open the door to your dark apartment, when suddenly the lights turn on and you hear a giant yell, "Happy Birthday!" You see 30 of your friends gathered in your home.</p>                              | <p>"Wow, what a surprise."</p>                             | <p>a. You are a bit annoyed you thought you made it clear to your roommate that you don't like surprise birthday parties.</p> <p>b. You are surprised by the fact that your friends managed to break in to your locked apartment without a key.</p> <p>c. You are excited and happy to see all of your friends.</p>   |
| <p>3. Group Project</p> <p>Your roommate asks you to look at a group project they've been working on for the past 3 weeks. You go to their room to have a look at it for the very first time.</p>  | <p>"So this is what you did? That's really something!"</p> | <p>a. You are impressed at your roommate's artistic and organizational ability.</p> <p>b. You see a horrific jumbled mess of a poster and cannot believe anyone would turn in something like that.</p> <p>c. You walk into a very cluttered room that has several "projects" in various stages of development and want to verify that you are looking at the correct one.</p>         |
| <p>4. Movie (Park Bench F'06)</p> <p>Your friend takes you to a movie on your birthday and pays for your ticket as a gift. You had been expecting a completely different kind of movie. After the film, your friend asks what you thought of it.</p> | <p>"It really wasn't what I expected."</p>                 | <p>a. You are quite satisfied with the movie, although you really would have preferred what you were expecting.</p> <p>b. You are pleasantly surprised by the unexpected type of movie and loved it.</p> <p>c. You hated it, but don't want to seem ungrateful.</p>   |