Visual signals and children's communication: negative effects on task outcome

Doherty-Sneddon, G, McAuley, S., Bruce, V., Langton, S.,

Department of Psychology, University of Stirling, Stirling, FK9 4LA

Blokland, A., and Anderson, A.H

Department of Psychology, University of Glasgow, Glasgow

Running Head: VISUAL SIGNALS AND CHILDREN'S COMMUNICATION

Abstract

Previous research has found that young children fail to adapt to audio-only interaction (e.g. Doherty-Sneddon & Kent, 1996), and perform difficult communication tasks better face-to-face. In this new study, children aged 6- and 10 year-olds were compared in face-to-face and audio-only interaction. A problem-solving communication task involving description of abstract stimuli was employed. When describing the abstract stimuli both groups of children showed evidence of face-to-face interference rather than facilitation. It is concluded that, contrary to previous research, for some communication tasks access to visual signals (such as facial expression and eye gaze) may hinder rather than help children's communication.

Children communicate using a number of different communication media, including face-to-face interaction, telephones, computers and even using video conferencing. These media differ from one another on a number of dimensions including bandwidth (information capacity of the telecommunications channel), the use of speech versus writing and the availability of visual communication signals from facial expressions, and facial and manual gestures. As children develop, they must learn to understand and to express information to others irrespective of the communication medium. The impact which visual signals have on children's communication is the focus of the current article.

Adults are skilled communicators who can adjust their communicative strategies to cope with different media and can often maintain their task performance regardless of the availability of visual signals (for example, Chapanis, Ochsman, Parrish & Weeks, 1972; Williams, 1977). However, even for adults, the presence or absence of visual signals does affect the process of communication. Boyle, Anderson and Newlands (1994) found that while the same level of task performance was maintained regardless of the availability of visual signals, significantly more words and turns were required for pairs of adult subjects to accomplish a problem-solving communication task when they could only hear each other compared with face-to-face interaction. There is considerable further evidence for the importance of the information transmitted by visual signals in the communication process (see also, McNeill, 1985; Clark & Brennan, 1991; Goldin-Meadow, Wein, & Change, 1992).

Children may be particularly dependent upon non-verbal signals in their communication attempts. Church and Goldin-Meadow (1986) found that when explaining their reasoning on conservation tasks, children transmit information via hand gestures which they do not verbalize. The authors suggest that when children attempt to convey difficult material in gesture form only, this demonstrates that at some level they understand the material but cannot yet verbalize it.

Doherty-Sneddon and Kent (1996) report that young children rely on visual communication to support their relatively poor language. They found a face-to-face performance benefit over unseen interaction for 4- and 6 year olds completing problem-solving communication tasks of the kind used in the Boyle et al (1994) study with adults. They conclude that when the information to be transmitted is demanding, visual signals, such as gesture, are central to a child's abilities. This makes sense if one accepts that gestures and speech have different complexities, and that young children process the less demanding information more readily (Feyereisen & de Lannoy, 1991; Church & Goldin-Meadow, 1986).

However, in the literature on adult communication, effects of visual signals are qualified by the communicative task. For example, while problem-solving task outcome is seldom influenced by the presence or absence of visual signals (e.g. Doherty-Sneddon, Anderson, O' Malley, Langton, Garrod & Bruce, 1997; Boyle et al, 1994), in more social tasks involving negotiation or conflict resolution, non-verbal signals can have positive effects on task outcome (Fish, Kraut, Root, & Rice, 1993; Short, Williams, & Christie, 1976). Visual signals may add value to more social tasks since they transmit emotional and social information which has an integral role to play in such tasks.

Furthermore, there is some evidence that visual signals may even interfere with the performance of some activities. For example, Glenberg, Schroeder, and Robertson (1998) report that when people are asked moderately difficult questions, they often avert their gaze. They demonstrate that the frequency of gaze aversion is related to the difficulty of cognitive processing, and that averting the gaze improves performance. They suggest that averting the gaze helps people to disengage from environmental stimulation and thereby enhances the efficiency of cognitive processing directed by non-environmental stimulation. Glenberg et al propose that while some tasks e.g. naming or object recognition tasks may be facilitated by environmental cues, conceptually driven tasks (and mental imaging is such a task) will be hindered. Other researchers have similarly suggested that gaze processing may increase cognitive load (Beattie, 1981; Ellyson, Dovidio, & Corson, 1981).

Tasks for which visual communication signals may interfere rather than facilitate are those which involve mental imaging. Ozols and Rourke (1985) propose a link between visuospatial processing problems and problems of processing visual-perceptual communicative information such as facial expressions and gestures etc., and suggest that the processing of visuospatial information and nonverbal communicative signals are related. It may be that these are linked by common neurological structures or perhaps by common underlying cognitive processes such as pattern recognition. Consistent with this are results described by Hanley, Young and Pearson (1991) reporting on patient ELD. From her pattern of deficits and abilities on visuospatial and verbal tasks, patient ELD illustrated the distinction between the visuospatial sketchpad (VSSP) and the phonological loop in working memory: ELD had deficits in VSSP while retaining an intact phonological loop. In addition to her impairment in VSSP, ELD was impaired on some tasks of face processing, suggesting a common substrate for social and non-social visuospatial processing. ELD was not impaired on all tasks of face processing however. Some aspects of expression recognition remained intact, although not necessarily all (Young, personal communication). In any case, demands of <u>dynamic</u> face processing may be greater than static. The online processing of dynamic facial expressions, eye gaze, gesture etc. may involve VSSP to a far greater extent than tests using static images.

The current study was designed to follow up previous work which has shown face-to-face <u>benefits</u> for communicative problem-solving tasks for adults and children (Doherty-Sneddon et al, 1997; Boyle et al, 1994; Doherty-Sneddon & Kent, 1996; McEwan, 1997). If visuo-spatial processing and the perception of visual communicative signals share common processing resources, then it follows that the presence of facial expressions and gestures may in some circumstances <u>interfere</u> with the processing of other kinds of visuospatial information.

In the new work reported here, children attempted to describe and to understand descriptions of complex, abstract shapes. Such a task requires that the information sender scrutinises the shapes for distinctive visual properties, and the information receiver must build a visual representation of the described shape over time, sufficient to select the correct target shape from distractors. These

materials differ from those with which we have previously found face-to-face benefits with young children (Doherty-Sneddon & Kent, 1996; McEwan, 1997). The earlier tasks, investigating the impact of different levels of visibility on children's communication, involved communication about pictures of real objects that were nameable to the children. For example Doherty-Sneddon and Kent used a Map Task (Boyle, et al, 1994) with 6- and 10-year old children. This task involves communication about a path around a schematic map containing features such as 'cows', 'houses' and 'trees'. We predict that when dealing with descriptions of abstract shapes the visuo-spatial demands for both partners may be such that visual signals from the face-to-face communicative context could interfere with performance on the task.

To investigate this, we studied pairs of six-year olds and pairs of ten-year olds performing the shape description task when they could see each other (but not the shape) and when they could only hear each other. Earlier work showed that 6-year olds often failed to adapt their communication strategies when they could not see one another compared with when they could, resulting in communicative performance deficits. In contrast, 10 year olds did not show performance deficits when they could not see one another while communicating (Doherty-Sneddon & Kent, 1996). We therefore chose these age groups in our current study to extend this work.

Method

Participants. The participants were 56, 6-year olds (mean = 6 years 5 months; range = 5 years 7 months- 7 years 2 months) and 64 10-year olds (mean = 10 years 3 months; range = 9 years 9 months - 11 years) from primary schools in the Stirling, Perthshire, and Clackmananshire areas. They were recruited through their schools following the acquisition of local authority, school and parental permissions.

Design, Materials and Procedure. Participants came to the University of Stirling for the experimental sessions in pairs and were collected by the experimenter in a departmental car. Each pair of children completed 3 shape description tasks (of increasing difficulty). The order of difficulty

6

always moved from easiest to hardest. Previous piloting with the tasks had confirmed that the materials were appropriately graded in terms of difficulty, and that the decreasing performance across the three tasks was due to this. Practice tasks were completed before the tasks proper. Each pair of children completed the shape description tasks in one visibility condition (face-to-face or audio-only), so this formed a between-subjects manipulation.

Visibility conditions within an age group were matched for numbers of single sex and mixed sex pairs (6 year-olds: 5 pairs of boys, 7 pairs of girls, 2 mixed sex in each condition; ten year-olds: 6 pairs of girls and 10 pairs of boys in each condition). The children were randomly assigned to the role of either instruction giver (describer of the shapes) or instruction follower (chooser of the shapes), and maintained the same role throughout.

The shape description task was a variation of the referential task designed by Glucksberg, Krauss and Weisberg (1966). The stimuli used were carefully designed to present the children with a task which would be communicatively challenging, but which they could achieve with some effort. Three versions of the task were used with increasing difficulty. Difficulty was increased by decreasing the 'nameability' or likelihood that the shape 'looked like' a real object such as a house (in other words, increasing the abstract nature of the stimuli). See Figure 1 for examples.

Figure 1 about here

The Instruction Giver was given a set of 5 blocks stacked in an opaque dispenser. The Instruction Follower had an array of 15 blocks in front of them occluded from the view of the Instruction Giver by a screen.

Each block had an individual design on one of its faces. The designs were chosen on the basis that the correct referent would not always be readily identifiable as there was potential for confusion between two or more blocks in the Instruction Follower's array. The different shapes and colours meant that children of this age would find describing them a fairly demanding but not impossible task. Five of the Instruction Follower's blocks matched exactly the designs on the 5 blocks which

the Instruction Giver possessed. The other blocks in the instruction follower array were distracters (2 distracters per target block) The task involved the Instruction Giver removing his/her blocks one at a time from the dispenser and describing them to the Instruction Follower so that he/she could choose the correct matching block from their referent array. The Instruction Follower then placed the chosen referent into their own stacking container so that the order of choices could later be checked by the experimenter. Task outcome was scored as the number of correct blocks chosen (out of five) in a particular trial. The distracter blocks were present partly to minimise the constraining of later responses by earlier ones. In fact 81% of dialogues contained only choices within the target and distracter group, and therefore responses were rarely constrained by a target having been erroneously chosen on a previous trial.

A low table was used and the children sat opposite one another. The table was sectioned in front of the children by a screen. In the face-to-face condition part of this was removed so that they could see one another.

The dialogues were audio recorded on a DAT (Sony DTC1000ES) using Shure SNIOA microphones. Transcriptions were made of the dialogues were produced from these recordings. Coding of Align and Check games from Conversation Games Analysis (Kowtko, Isard, & Doherty 1991) was performed on the dialogues.

Conversational Games Analysis. This analysis bears a family resemblance to the models proposed by Sinclair and Coulthard (1975) and Traum and Hinkelman (1992). These kind of models are useful tools in analysing the pragmatic functions of utterances with respect to achieving speakers' goals. The analysis charts the way speakers achieve their communicative goals. Conversational Game theory proposes that the achievement of the goals and subgoals of conversation occur through the accomplishment of dialogue units called Conversational Games (Power, 1979; Houghton & Isard, 1987). The term "game" here is an analogy which is used to capture the fact that conversational units have rules which both participants know and follow; they have a beginning and

8

an end, and they are interactive. That is, a game can only be accomplished through the interaction between two or more participants.

For example, if a speaker wishes to ask another person to carry out some action, he or she can use an **INSTRUCT** Game to do so. This would be initiated by an utterance which states the instructional goal (e.g., "please pass me the salt"). If this is a reasonable request which is within the capabilities of the listener then the expectation would be that the listener will carry out the required action. If this is so, then the **INSTRUCT** Game would come to an end here. However, more extensive interaction may be required in order to provide clarification to the listener and this may result in questions being embedded within the original instruction (e.g., "did you say the salt or the sauce?").

In the original version of the analysis system (Kowtko et al, 1991) there are 6 types of conversational game. Doherty-Sneddon et al investigated how adults adapt to audio-only interaction using Conversational Games Analysis. They found that adults adjusted their dialogue structure in response to audio-only interaction by increasing the number of times they monitored their listener's comprehension (Align games) and by increasing the number of times they checked their own understanding of a previous message (Check games). Doherty-Sneddon et al propose that such changes in dialogue structure reflect the role that visual cues play in delivering feedback information. The following example illustrates how align and check games might be used within everyday conversation. Speaker A is describing the location of her house to speaker B. Underlined sections are the initiations of the align game and check game.

Speaker A: Once you've come through the main part of the village, you'll see a church on your right. You want to turn left there. <u>Your know where I mean?</u> **(Begin align game)**

Speaker B: <u>Hang on. Do you mean turn left at the road directly opposite the church?</u> **(Begin check game)**

Speaker A: Yes that's the one. **(End check game)**

Speaker B: Okay. **(End align game)**

Given the particular importance of align and check games in adults' adaptation to audio-only communication, the occurrence of these game types was investigated in the children's dialogues of the current study. The dialogues from the easiest and hardest shape description tasks were coded for Align and Check games (128 dialogues in total) to investigate whether the children adapted to audio-only interaction in the ways in the same way that adults did.

An interjudge reliability was performed on a random 11% sample of these dialogues. The independent coder agreed on 92.6% of the original classifications.

Results

Task outcome. Table 1 shows the mean scores (out of 5) in each condition of the experiment. A 3-way mixed ANOVA was used to analyse these task performance scores. Visibility condition and age were between-pair variables (2 levels: seen and unseen; 6-years and 10-years respectively), and task difficulty was a within-pair variable (3 levels: Shape description task 1, 2, and 3). Visibility condition had a significant effect on task performance, $F(1,56) = 7.32$, $p<.01$, with unseen performance scores higher than face-to-face scores. Age also had a significant effect, $F (1,56) = 18.56$, $p<.01$, with older children performing better than the younger children. Increased task difficulty produced a significant decrease in performance, $F(2,112) = 64.40$, $p<.001$ (mean task 1= 4.04; mean task 2= 3.80; mean task 3= 2.28). Planned comparisons t-tests revealed that Shape description tasks 1 and 2 did not differ from one another but did both differ from the hardest version of the task ($t(112) = 6.39$, $p<.01$; and $t(112) = 7.39$, $p<.01$ respectively). No interactions were found.

Tables 1 and 2

Dialogue length. The change in task outcome across visibility conditions is accompanied by an increase in words in the unseen condition. Means are given in table 2. Children use more words in this task when they cannot see each other, and this appears to help their performance. A 4-way ANOVA carried out on the number of words gave a significant effect of visibility condition ($F(1,112) = 10.29$, $p<.01$), with unseen dialogues longer than face-to-face ones (mean face-to-face = 81.4 words; audio-only = 113.2 words per dialogue). Task role also had a significant effect, F

(1,112) = 27.81, p<.001, with instruction givers saying more than instruction followers (mean instruction giver =124.8 words; mean instruction follower = 69.88). Task difficulty had a significant effect, F (2, 224) = 57.72, p<.001, with the more difficult task dialogues longer than the easier. This effect interacted with visibility (F (2,224) = 7.26, p< .01). This shows that the increase in dialogue length associated with audio-only interaction is particularly strong for the hardest shape description task, illustrating the increase in verbal effort required for this task. Simple effects analyses revealed a significant increase in dialogue length with increasing task difficulty in the audio-only condition, F (1,112) = 8.77, p<.01, but a non-signficant trend for the same effect in the face-to-face condition, F (1,112) = 3.6, p = .06. Children appear to be showing better adaptation to the increased task demands of the harder trials in the audio condition. Finally, the effect of task difficulty interacted with task role, F (2,224) = 4.57, p<.05. Simple effects analysis revealed that it was only instruction givers who significantly increased their verbal effort with increasing task difficulty, F (1,112) = 5.41, p<.05.

Both groups of children therefore have longer audio-only dialogues which result in better task performance. It appears that not having access to visual signals allows them to adapt their communication better to the task demands. The following extracts show two different descriptions of the same block. The audio-only description is more elaborate and results in a correct choice. The seen description is inadequate and results in an incorrect choice. The more elaborate description is very much mediated by the instruction follower's repair procedures. These are underlined and are coded as check games in the conversational games analysis, the results of which are reported below. These are analogous to the requests for clarification reported by Lloyd (1992) as being central to the success of audio-only referential communication of children (10 year olds) and adults.

Audio-only description:

Instruction Giver:	would...now this ones got two red triangles none shiny none shiny/
Instruction Follower:	are they wee?
Instruction Giver:	one wee and one big...and there's a wee bit in the middle whats whats just left and/
Instruction Follower:	is there a gap?
Instruction Giver:	a gap yeah there's a gap

Face-to-face shape description:

Instruction giver: right its something that's got red triangle two red triangles...there is so blue

In order to investigate whether the effect of visibility was mediated by longer descriptions in the audio-only dialogues regression analyses were carried out. The dependent variable was the average task performance across the three shape description tasks. The predictors were visibility condition and average dialogue length (in number of words). When dialogue length was the first predictor it accounted for an amount of the variance in task performance that approached significance, $F (1,58) = 2.92$, p=.09. The joint predictive value when visibility condition was added was significant, $F (2,57) = 3.40$, p<.05. When visibility condition was the first predictor it accounted for a significant amount of the variance, $F (1,58) = 5.60$, p<.05. When dialogue length was added this did not add to the predictive power, $F (2,57)= 3.40$, P<.05 (variance associated only with dialogue length was not significant).

The improvement in performance in the audio-only dialogues does not therefore appear to be mediated only by an increase in the amount that is said by the children. It was therefore important to look more qualitatively at the strategies for communication that the children were using to see whether these differed across visibility condition. Conversational games analysis provides one way of doing this.

Conversational Games Analysis. A 3-way anova was conducted with the number of align games initiated by instruction givers in each dialogue the dependent variable (only 1 align game in the corpus was initiated by an instruction follower- task role was therefore not included as a variable in this analysis). Task difficulty (2 levels) was a within-subject variable. Age and visibility condition were between subject variables. Task difficulty and visibility condition influenced the number of align games initiated in each dialogue. Means are given in table 3. Visibility condition had a significant effect on the occurrence of align games, $F (1,56) = 5.95$, p < .05, with more align games being used in the audio-only interactions. This is a similar pattern to that found with adults

(Doherty-Sneddon et al, 1997). Task difficulty had a significant effect $F(1,56) = 6.30$, $p<.05$ with more align games initiated in the harder shape description task.

A 4-way anova was conducted with the number of check games initiated by each participant the dependent variable. Task difficulty was a within subjects variable. Age, visibility condition and task role were between subject variables. Visibility, task difficulty and task role all influenced the occurrence of check games. Means are given in table 4. Significantly more check games were produced in unseen interactions, $F(1,112) = 4.0$, $p<.05$. In addition instruction followers produced more check games than instruction givers, $F(1,112) = 54.55$, $p<.001$. Finally more check games were produced with the harder version of the task, $F(1,112) = 28.9$, $p<.001$. While there was no main effect of age, there was an interaction between task difficulty and age, $F(1,112) = 4.28$, $p<.05$. This showed that the older children initiated more check games than the younger ones only in the harder task. There were also significant interactions between task difficulty and task role, and task difficulty and visibility condition ($F(1,112) = 24.63$, $p<.001$; $F(1,112) = 7.96$, $p<.01$ respectively), with the effect of task difficulty being mediated primarily by instruction followers, and the effect of visibility occurring in the most difficult version of the task.

Gaze analysis. An analysis of the gazing behaviour of the children was carried out to establish whether or not the children actually looked at one another and to what extent. The 16 face-to-face dialogues in the 10-year olds sample and 11 face-to-face dialogues in the 6-year olds' sample were coded for gazing behaviour (video data for 3 of the 6-year olds dialogue was lost due to a technical fault). Two independent coders marked the incidences of eye gaze on transcripts of the dialogues. A random sample of 22% of the dialogues was coded by both and an interjudge reliability of 83% was found. Overall the amount of gaze while speaking and listening for the children was 20% and 24% of words (spent in either speaking or listening respectively).

There was a trend for the younger children to decrease the amount that they gazed in the harder version of the task, $F(2,40) = 2.81$, $p =.07$ (mean percentage of words = 28%, 26%, and 20%). With the older children the only change across conditions was that the instruction followers

decreased the amount that they gazed at their partners while listening as the task became harder (means = 14.8%, 13.5%, and 9.5% respectively).

Pearsons correlations were carried out between the task scores for each level of task difficulty and the percentage of words associated with gaze. Task score and gaze level did not correlate consistently. For the younger children task score and speaker gaze correlated positively in the easiest and hardest versions of the task (r(11) = .48, p= .07; and r(11) = .677, p<.05 respectively). A similar pattern was found for the younger children's listener gaze (r(11) = .51, p=.06; r(11) = .48, p =.07). For the older children the only correlation that approached significance was between speaker gaze and task score, r(16) = .42, p=. 051. Therefore contrary to our expectations (based on our visual interference hypothesis) increasing levels of gaze were related to increasing task performance.

Discussion

For children of these ages the shape description task is performed more effectively when they cannot see one another, and this improved performance is accompanied by longer dialogues, although the increased length of dialogues does not seem in itself to mediate the improved performance. The performance benefit for the audio condition is in contrast with a number of previous studies using different kinds of communicative task where children do better in face-to-face interaction (e.g. Doherty-Sneddon & Kent, 1996). In addition, Doherty-Sneddon and Kent (1996) report that preschool children produced longer audio-only dialogues compared to face-to-face interactions and face-to-face performances were better. This shows that increased verbal effort is not always effectively translated into improved task performance. The current results suggest that visual social cues may be disruptive for certain tasks. The unseen dialogues were longer, suggesting that in this condition the children put more effort into their speech. The regression analyses performed with dialogue length and visibility as predictors of task performance revealed that it was not just the increased amount of speech that mediated improvement when the children could not see one another.

We propose that the children say more in the audio-only condition because they concentrate better on the task when they cannot see one another. We propose two alternative explanations for this. First, visual signals in face-to-face interaction may overload something like the visuospatial sketchpad thereby interfering with children's abilities to process (either to express or to understand) the required visuospatial information. Alternatively, it may be that in the face-to-face interactions children used a style of interaction that relied upon visual communication signals- this is typically found in children's face-to-face interactions (Doherty-Sneddon & Kent , 1996). Furthermore, this face-to-face style contrasts with an audio-only style (involving increased verbal checking and aligning of mutual knowledge), that fits better with task demands.

The current task does not benefit from the use of e.g. gesture in the way that the map task does, where for example, hand gestures can be beneficial in describing the shape of the route. Success on the current task relies more heavily on fine tuning the verbal establishing of mutual understanding in order to chose a shape that is difficult to name from a relatively large array of potential referents. This contrasts with the map task where not only are the features on the map nameable objects, but where any given name primarily refers to only one location on the map (sometimes two). Therefore while the map task involves communication about visuospatial information (the shape of the route), visual communication signals are likely to be of benefit and to complement the encoding and decoding of verbal information[1] (see for example, McNeill, 1985). In contrast visual communication signals in the shape description task are unlikely to be informative and are more likely to interfere with the processing of task related visuospatial information. In other words, face-to-face communication has costs that lead to an increase in cognitive load (as proposed by e.g. Beattie, 1981; Glenberg et al, 1998). For tasks that benefit from e.g. gesture and posture (such as the map task) there will also be communicative benefits associated with face-to-face interaction that outweigh the costs. This hypothesis fits well with information processing accounts of communication. For example, Shatz (1985) suggests that the deployment of communication skills depends partly on the processing demands of the task- that if task demands are high that skills

---

[1] The children in the current study also did another task- the picture placement task. This task is similar to the map task in that it involves the description of nameable pictures and their location on a board relative to one another. The

children possess may not be utilised. The cognitive load of the current face-to-face interactions that may have interfered with their abilities to implement verbal communication skills such as check and align games that would have facilitated the establishing of mutual understanding.

In support of the first, Glenberg et al (1998) report that averting gaze helps people (adults) to disengage from environmental stimulation and thereby enhances the efficiency of cognitive processing. Glenberg et al make the distinction between tasks involving naming or object recognition and conceptually driven tasks, with the latter prone to interference from environmental stimulation. The shape description task we used in the current study involves a conceptually driven component since the stimuli are abstract and therefore not nameable, and this is the kind of task for which Glenberg et al would suggest environmental stimulation could disrupt at a general level. In addition, however, the processing of visual communicative signals might compete directly with mental imaging for a finite, short-term visuospatial processing capacity (e.g. the visuospatial sketchpad in Baddely's 1986 model of working memory). In these circumstances, audio-only interaction may allow existing visuospatial abilities to be used exclusively for the production of visual representations of the stimuli used in the task. This proposal is consistent with other suggestions of a common substrate for social and non-social visuospatial processing (for example: Ozols & Rourke, 1985; Hanley, et al, 1991).

In support of the second explanation is work done on adult adaptation of speech intelligibility in response to the presence or absence of visual communication. Anderson, Bard, Sotillo, Doherty-Sneddon and Newlands (1997) found that adults reduced the intelligibility of their speech in face-to-face interaction compared with audio-only. This might suggest that when visual signals (such as lip configuration and facial expression) are available less effort is required in articulation. However although speakers were found to speak less clearly when they could see one another there was no evidence that they tracked their listeners' use of visual information or that they articulated less clearly only when they observed the listener exploiting such information. The authors suggest that speakers assume that listeners are viewing them in face-to-face interaction (and therefore benefiting

---

younger children were at floor on this (hence why the data is not reported here), however the older children did

from visual information). In audio-only interaction speakers articulate more clearly since they know that their listeners cannot take advantage of visual information. In the current study it may be that the children assume that they understand one another better when they can see one another when in fact they do not. Indeed attention to the task materials may detract from important monitoring of feedback information from interlocutors. In contrast, the children take less for granted when they cannot see one another and put more effort into establishing mutual understanding.

For whatever reason, the children accomplished the shape description task better when they couldn't see one another. We were therefore interested in the ways in which their audio-only dialogues differed from their face-to-face ones. Previous work has found that adults adapt to audio-only interaction in particular ways (e.g. Boyle et al, 1994, Doherty-Sneddon et al, 1997). In the current study we wanted to investigate whether any adaptation was evident in the children's dialogues which might have contributed to their improved performance. It was found that the instruction givers used significantly more align games in the unseen interaction. In other words instruction givers attempted to make sure that their instruction followers had understood them more frequently when they could not see them The following extract illustrates the use of an align game in a 10 -year olds dialogue. The initiation of the align game is underlined.

Instruction giver: Right, the next one's a pentagon ... no a hexagon with like a wee triangle cut out a red hexagon with a wee ... triangle cut out.   <u>You got it? ... You got it?</u>
Instruction follower: No I don't have any hexagons ... or triangles ... is it ... is it a pentagon?
Instruction giver: Oh no, it's a pentagon.

Furthermore instruction followers checked their own understanding in the audio-only context more often than in the face-to-face condition. Both these results show that the children adapt their interactional style in audio-only interaction. Such adaptations may partially explain their improved performance in the unseen condition since it may have facilitated the successful establishment of mutual understanding. In support of this, Doherty-Sneddon (1995) reports a correlation between the

---

significantly better in face-to-face interaction compared with face-to-face interaction.

number of check games initiated by adults in a problem-solving task and improved task performance. This is consistent with the findings of Lloyd (1992) showing that poor instructions can be compensated for by the message receiver's ability to ask appropriate questions. In other words, audio-only interaction increases the likelihood that children will explicitly check their own and each other's understanding. This in turn provides valuable opportunities to resolve communicative difficulties.

In addition, Doherty-Sneddon (1995) and Anderson, O'Malley, Doherty-Sneddon, Langton, Newlands, Mullin, Fleming and Van der Velden (1997) report that eye gaze often seems to occur in face-to-face interactions at the same points in dialogue structure as additional verbal alignment and checking occurs in audio-only dialogues. These authors suggest that align and check games in audio-only interaction may substitute for visual strategies in face-to-face interaction (such as looking at the facial expression of an interlocutor). If the children are attending to their task materials then access to visual feedback information might be affected, and they may presume that they fully understand one another when they do not.

The question remains whether the changes in verbal strategy reflect an adaptation to the loss of visual communication signals (as previous research has suggested e.g. Doherty-Sneddon and Kent, 1996). Alternatively it may show a focusing on verbal communication strategies when distracting visual information is reduced. This interpretation is supportive of the Glenberg et al (1998) proposal that disengagement from extraneous environmental information can facilitate cognitive processing.

An analysis of the gazing behaviour of the children was carried out to establish whether or not the children actually looked at one another and to what extent. Overall the amount of gaze while speaking and listening for the children was 20% and 24% of words (spent in either speaking or listening respectively). Argyle (1990) reports that adults typically spend about 40% of speaking time engaged in gaze and 60% of listening time. However these figures come from studies of non-structured conversations. When there is an object of legitimate attention gazing frequency can decrease dramatically, to as little as 6.4% (Argyle and Graham, 1977). Given the task-oriented

nature of the shape description task the levels of gazing by the children are not low in comparison with expected levels of adult gaze. The children therefore do look at one another while doing the shape description task in face-to-face interaction. This suggests that the decline in face-to-face performance is not solely due to the children simply missing important visual information.

In addition, the gaze analyses provided some support for our prediction that visual signals accessed while gazing might produce an increased cognitive load for this task. There was a trend for the younger children to decrease the amount that they gazed in the harder version of the task. With the older children the only change across conditions was that the instruction followers decreased the amount that they gazed at their partners while listening as the task became harder (means = 14.8%, 13.5%, and 9.5% respectively). These findings provide some support for the prediction that gaze aversion will increase when tasks become harder in order to decrease the cognitive load associated with gaze (Glenberg et al, 1998).

However if our hypothesis is correct and visual signals interfere with the visuospatial demands of the shape description task we would expect a negative correlation between the frequency of gazing and task success. Pearsons correlations were carried out between the task scores for each level of task difficulty and the percentage of words associated with gaze. Task score and gaze level did not correlate consistently. For the younger children task score and both speaker and listener gaze correlated positively in the easiest and hardest versions of the task. For the older children the only correlation that approached significance was between speaker gaze and task score. Therefore contrary to our expectations (based on our visual interference hypothesis) increasing levels of gaze were related to increasing task performance. This suggests that the children obtained some useful information when they gazed at their partners in face-to-face interaction. These results support the idea that the problem with the face-to-face interactions is that the visual access brings about a communicative style that is less conducive to task success.

These findings have potentially important implications. There may be certain communicative situations where social signals may interfere with other cognitive activities, which implies that

certain kinds of activities – e.g. the learning of geometrical concepts, may be more difficult if conducted in an environment rich in social cues. This does not negate the potentially important role that visual social signals serve in many other types of communication as discussed by Doherty-Sneddon and Kent (1996), but suggest that we must not assume that visual signals will always be beneficial.

References

Anderson, A.H., Bard, E., Sotillo, C., Doherty-Sneddon, G., & Newlands, A. (1997). The effects of face-to-face communication on the intelligibility of speech. <u>Perception and Psychophysics,</u> <u>59</u>, 580-592.

Anderson, A.H., O'Malley, C., Doherty-Sneddon, G. Langton, S., Newlands, A., Mullen, J., Fleming, A. & Van der Velden, J. (1997). The impact of VMC on collaborative problem solving: an analysis of task performance, communicative process and user satisfaction. In A.J. Sellen, Finn & S. Wilbur (Eds) <u>Video Mediated Communication</u>. LEA.

Argyle, M. (1990). Bodily Communication. Routledge: London.

Baddeley, A. W. (1986). <u>Working Memory</u>. Oxford: Oxford Univeristy Press.

Beattie, G.W. (1981). The Regulation of Speaker Turns in Face-to-Face Conversation: Some Implications for Conversation <u>Semiotica</u>, <u>34</u>, 55-70.

Boyle, E.A., Anderson, A.H., & Newlands, A. (1994). The Effects of Visibility on Dialogue and Performance in a Cooperative Problem Solving Task. <u>Language and Speech, 37 </u>(1), 1-20.

Chapanis, A., Ochsman, R.B., Parrish, R.N., & Weeks, G.D. (1972). Studies in interactive communication: the effects of four communication modes on the behaviour of teams during co-operative problem-solving. <u>Human Factors, 14</u>, 487-509.

Church, R.B. & Goldin-Meadow, S. (1986). The Mismatch between Gestures and Speech as an Index of Transitional Knowledge. <u>Cognition, 23</u>, 43-71.

Clark, H.H, & Brennan, S.E. (1991). Grounding in Communication. In L.B. Resnick, J. Levine, & S.D. Teasley (Ed's), <u>Perspectives on Socially Shared Cognition</u>.Washington, D.C.: American Psychological Association.

Doherty-Sneddon, G. (1995). <u>The Development of Conversational and Communication Skills</u>. Unpublished PhD Thesis, University of Glasgow.

Doherty-Sneddon, G. & Kent, G. (1996). Visual signals in children's communication. <u>Journal of Child Psychology and Psychiatry</u>, 37, 949-959.

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S. & Bruce, V. (1997). Face-to-face and video-mediated communication: a comparison of dialogue structure and task performance. <u>Journal of Experimental Psychology: Applied</u>, 3, 105-125.

Ellyson, S.L., Dovidio, J.F., & Corson, R.L. (1981). Visual differences in females as a function of self-perceived expertise. <u>Journal of NonverbalBehaviour</u>, <u>5</u>, 164-171.

Feyereisen, P., and de Lannoy J.D. (1991). <u>Gesture and Speech</u>. Cambridge University Press.

Fish, R.S., Kraut, R.E., Root, R.W, & Rice R. (1993). Video asTechnology for Informal Communication. <u>Communications of ACM, 36</u>, 48-61.

Glenberg, A.M., Schroeder, J.L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. <u>Memory and Cognition</u>, <u>26</u>, 651-658.

Glucksberg, S., Krauss, R.M., & Weisberg, R. (1966). Referential Communication in Nursery School Children: Method and Some Preliminary Findings. <u>Journal of Experimental Child Psychology</u>, <u>3</u>, 333-342.

Goldin-Meadow, S. Wein, D. & Chang, C (1992). Assessing Knowledge Through Gesture: Using

Children's Hands to Read Their Minds. Cognition And Instruction, 9(3), 201-219.

Hanley, J.R., Young, A.W., & Pearson, N.A. (1991). Impairment of the visuo-spatial sketch pad. Quarterly Journal of Experimental Psychology Human Experimental Psychology, 43A(1), 101-125

Houghton, G., & Isard, S. (1987). Why to speak, what to say, and how to say it. In P. Morris (Ed) Modelling Cognition, Wiley, (pp.249-267).

Kowtko, J., Isard, S., & Doherty, G. (1991). Conversational Games in Dialogue. Technical report HCRC publications University of Edinburgh. A. Lescardes (ed).

Lloyd, P. (1992). The role of clarification requests in children's communication of route directions by telephone, Discourse Processes, 15, 357-374.

McEwan, R. (1997). Communication Across Different Media: The Importance of Seeing as well as Hearing for Preschool Children. Unpublished M.Sc. Thesis, University of Stirling.

McNeill, D. (1985). So You Think Gestures are Non-verbal? Psychological Review, 92, 350-371.

Ozols, E.J., & Rourke,B.P. (1985). Dimensions of social sensitivity in two types of learning-disabled children. In Neuropsychology of learning disabilities: Essentials of subtype analysis. (Byron P. Rourke, Ed.), pp. 281-301. Guilford Press, New York, NY, US

Power, R. (1979). The Organisation of Purposeful Dialogues. Linguistics, 17, 107-152.

Short, J., Williams, E., & Christie, B. (1976). The Social Psychology of Telecommunications. London.Wiley.

Sinclair, J.M. & Coulthard, R.M. (1975). Towards an analysis of discourse: The English used by teachers and their pupils. London: Oxford University Press.

Williams, E. (1977). Experimental Comparisons of Face-to-Face and Video Mediated

Communication. <u>Psychological Bulletin, 84</u>, 963-976.

Table 1. Mean shape description performance scores (maximum score = 5) in face-to-face and audio-only interaction

	Shape description 1	Shape description 2	Shape description 3
6 years face-to- face	3.43	2.71	1.57
6 years audio only	3.71	3.71	2.28
10 years face-to-face	4.19	4.19	2.38
10 years audio-only	4.81	4.56	2.88

Table 2. Mean number of words spoken in shape description task in face-to-face and audio-only interaction

	Shape description 1	Shape description 2	Shape description 3
Face-to-face: 6 years	54.0	89.8	88.4
Audio-only: 6 years	79.5	117.6	176.1
Face-to-face: 10 years	62.1	79.7	119.5
Audio-only: 10 years	68.2	94.0	150.8

Table 3. Mean number of Align Games in face-to-face and audio-only dialogues

	Shape description 1	Shape description 2
Face-to-face	.42	.77
Audio-only	1.31	2.0

Table 4. Mean number of Check Games in face-to-face and audio-only dialogues.

	Shape description 1	Shape description 3
Instruction giver: face-to-face	.03	.13
Instruction giver: audio-only	.07	.24
Instruction follower: face-to-face	2.33	3.87
Instruction follower: audio-only	2.77	7.63

Figure caption.

Figure 1. A set of blocks used in the shape description task. The far right column are the instruction giver's blocks, the other are the instruction follower's.

Author Notes