Interactional Context in Graphical Communication

Patrick G. T. Healey
Information, Media and Communication Research Group
Department of Computer Science,
Queen Mary University of London.

Simon Garrod, Nicholas Fay
HCRC
Department of Psychology,
University of Glasgow.

John Lee and Jon Oberlander
HCRC and Department of Architecture,
Division of Informatics,
University of Edinburgh.

Abstract

A body of empirical evidence indicates that interactional context has a key influence on the form and interpretation of language. This paper reviews a series of experiments which indicate that interactional context also plays a key role in the interpretation of drawings and sketches. Two graphical communication tasks, analogous to definite reference tasks, are described. The findings from these tasks show significant parallels between the mechanisms of co-ordination in graphical dialogue and natural language dialogue. Specifically; the coordination of graphical representation types by 'dialogue' participants, the contraction of recurrent 'graphical referring expressions', effects of direct interaction on the use of abstract drawings, and the development of community-specific graphical conventions.

Interactional Context in Dialogue

Conversation is a, if not *the*, key context of understanding for language. People's use of language to represent objects, events and situations is sensitive to, amongst other things; who they are speaking to, the mutual availability of referents, the history of their conversation and their (dis)joint membership of cultural and linguistic sub-communities (Hymes, 1972; Clark, 1998). Evidence for the direct influence of interactional context on interpretation and understanding comes from a variety of sources (see Krauss and Fussell, 1996, for a review). One example is provided by work on the Collaborative Model of dialogue (Schober and Clark, 1989; Clark and Wilkes-Gibbs, 1986). Wilkes-Gibbs and Clark (1992) have shown that full understanding of referring expressions depends on the degree of active participation in conversation by speaker and addressees. Non-active participants in a conversation, such as passive side-participants, overhearers, or bystanders, have more difficulty in interpreting referring expressions than active participants. This is observed even when, in gross informational terms, they are equivalent to active participants.

A second example of the influence of interactional context comes from studies of conceptual and linguistic co-ordination in dialogue. Garrod and Anderson (1987) have shown that conversational partners tend to match or 'entrain' on the form and interpretation of utterances during interaction. Where several types of semantically distinct referring expressions are possible for describing a location, people show a strong preference for matching the type of expression used by their conversational partner. Brannigan, Pickering and Cleland (2000) have observed similar entrainment effects with syntax. Garrod and Anderson (1987) argue that these dialogue phenomena reflect the operation of a basic dialogue co-ordination mechanism which simplifies the processes of production and comprehension in interaction.

Interactional Context in Graphical Dialogue

Intuitively, it might be supposed that graphical representations would be less sensitive to interactional context. One reason for this is that the production and use of drawings and sketches is normally treated, and analysed, as an activity more akin to monologue than dialogue (cf. Scaife and Rogers, 1996). There is evidence, however, that this underestimates both the actual and potential use of drawing activities as a mode of interaction. Anecdotally, drawings are often incrementally produced and modified as part of a conversational exchange. For example, sketch maps and explanatory diagrams form a familiar extension of many routine conversations.

van Sommers (1984) provides evidence from a questionnaire study that approximately half of routine, non-work, drawing activities take place with or for an audience. Although van Sommers does not report how often these interactions involve direct graphical exchanges, his findings demonstrate the variety of interactional contexts in which drawing occurs. The most frequently cited category is the production of sketch maps of a local area, either as part of an explanation or in order to give directions. The second most frequently cited category relates to activities with children including; games and amusements, teaching or helping with homework and helping children learn to draw. Additional categories of 'public' drawing include; sketching of hair, makeup and clothing, sketching house plans, drawing to express feelings, defacing pictures and drawing people.

The collaborative development and modification of sketches is a feature of many specialised work related interactions, such as architect-architect and architect-client (Neilson and Lee, 1994). We estimate that in the architects' practice studied by Healey and Peters (2001) approximately 30% of daily drawing activities occurred as an integrated part of a conversational exchange. Engle (1998) provides experimental evidence that graphics, gesture and language combine in explanatory dialogues to create composite communicative signals (cf. Clark, 1996). Overall, there is a *prima facie* case that sketches and drawings are often closely integrated into interaction and that this may have significant implications for their interpretation.

A second source of scepticism about the role of interactional context in the interpretation of sketches and drawings is the intuition that drawings and sketches are easier to interpret than language. Arguably, many of the interactional influences on language interpretation are associated with the conventional nature of linguistic representation. Coordinated interpretation of utterances requires the concerted application of conventions. Interaction is used to maintain and modify those interpretations. Drawings and sketches can exploit iconicity to provide a less arbitrary form of representation. Consequently, we might suppose that they would be less dependent on interaction to secure their interpretation. This idea is most plausible for, say, sketches of buildings or people but it does not cover the range of uses to which sketches and drawings are put. Explanations involving sketches of Venn diagrams or Euler circles provide perhaps the most obvious counterexample.

Experiments on Graphical Dialogue

The present paper summarises the findings from a series of experiments which, considered together, provide evidence that the interpretation of drawings and sketches is sensitive to interactional context. In particular, that interactional context affects the form, interpretation and understanding of sketches; and that the mechanisms and processes that give rise to these effects show important parallels to those identified for natural language dialogue.

The findings reported below are drawn from experiments involving two basic referential communication tasks, the Concept Drawing Task and the Music Drawing Task, in which pairs of subjects communicate about a variety of concepts using exclusively graphical means. These tasks can be thought of as two-way or conversational variants of the party game 'Pictionary' (TM).

The Experimental Tasks

The basic Concept Drawing Task uses an ordered list of twelve concept words drawn from the categories; places (e.g., "theatre", "art gallery", "museum"), people (e.g., "Robert de Niro", "Arnold Schwarzenneger", "Clint Eastwood"), television programmes (e.g., "drama", "soap-opera", "cartoon"), objects (e.g., "television", "computer" "microwave"), and abstract concepts (e.g., "loud", "homesick", "poverty"). One participant, the 'Drawer', is given an ordered list of twelve words. Their partner, the 'Chooser', is presented with an unordered list of the same twelve words plus four distractors. The task is for the Drawer to take each word in turn from their list and produce a sketch of it so that their partner, the Chooser, can identify the concept depicted. The aim is for the Chooser to determine the original ordered list of twelve concept words that the Drawer started with.

The basic Music Drawing Task is similar to the Concept Drawing task but uses pieces of music in place of concept words. The pieces are relatively unknown 30 second piano solos in a variety of genres and styles. In the typical procedure, the Drawer and the Chooser are seated in separate rooms. The Drawer listens to a target piece of piano music and produces a sketch of it. The Chooser has two pieces of music, the target and a distractor, and tries to select which piece is the one depicted by the Drawer. Playback of the pieces is self-paced and all drawing takes place on a shared virtual whiteboard which logs the drawing data for analysis (Healey, Swoboda, King, forthcoming).

In both tasks, subjects are free to draw anything they like; the only restriction is that they do not use letters or numbers. The types of drawing produced for each concept or piece of music varies substantially between pairs, some examples are provided in Figures , 2 and 3. All things being equal, each pair tends to establish their own conventional solutions to the communication problems posed by the task. Subjects appear to find both tasks enjoyable and engaging and perform them with above chance accuracy.

Effects of Interactional Context

A number of experiments have been performed using these tasks which suggest important parallels between the effects of interactional context on graphical and verbal dialogue. Here we provide an overview of the findings from these experiments and discuss their implications for investigations of graphical representation and models of human interaction.

Interactional Entrainment. One of the simplest pieces of evidence for effects of interactional context on the use of drawing comes from the Music Drawing task. Participants in this task produce drawings that can be reliably classified into two basic types[1]; 'Abstract' and 'Figurative' (Kappa = 0.9, N = 287, k= 2). Abstract drawings, illustrated in Figure 1,

[1]For ease of exposition a third, 'Composite, type is not discussed here

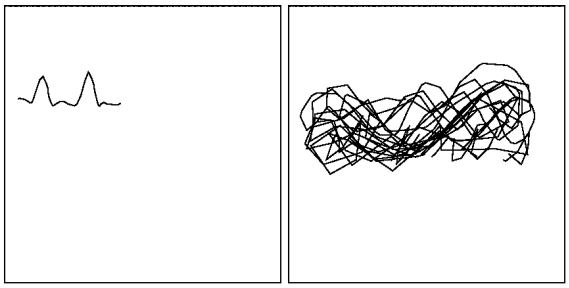

Figure 1: Example Abstract Drawings from Two Trials of the Music Drawing Task

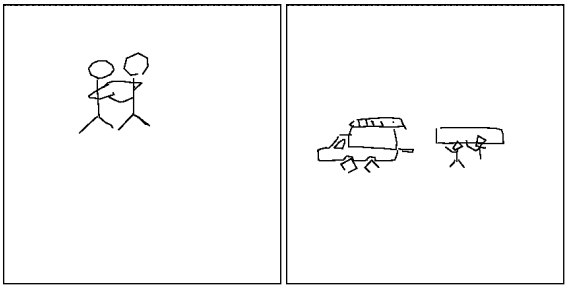

Figure 2: Example Figurative Drawings from Two Successive Trials of the Music Task

typically involve graph-like representations of e.g., pitch, melody, rhythm or intensity. By contrast, Figurative drawings, illustrated in Figure 2 typically depict recognisable objects, figures or scenes. Where pairs of participants in the task both take the role of Drawer (either by alternating roles or in manipulations in which both participants draw at the same time) they show a reliable tendency to match each another in their use of drawing the Figurative and Abstract drawing types (Healey, Swoboda, Umata, & Katagiri, 2001). As noted above, this pattern of entrainment between the participants in an interaction is also established for semantic and syntactic aspects of utterances in dialogue (Garrod and Anderson, 1987; Brannigan, Pickering and Cleland, 2000). Garrod and Anderson (1987) argue that entrainment constitutes a basic mechanism through which conceptual co-ordination is achieved in dialogue.

Contraction of Recurrent References. The procedure for the Concept Drawing task typically requires pairs to repeat the same set of twelve target words, in different orders, over several trials. This manipulation ensures that each word is drawn, and identified, several times by each pair. This is designed to reproduce the procedure followed by Clark and Wilkes-Gibbs (1986) who investigated the production of recurrent (verbal) referring expressions by conversational partners. Clark and Wilkes-Gibbs found that both the average number of words and average number of turns used to refer to a target item (in their case a tangram figure) rapidly declined with the number of repetitions. Experiments with the concept drawing task show the same pattern of reduction. This is illustrated in Figure 3 which shows a sequence of four trials (ordered left to right and top to bottom). Where target concepts recur, the drawings that represent them quickly become simplified. This is indexed both by simple quantitative measures such as the amount of 'ink' and number of lines used, and by their complexity as estimated by human judges and calculations of their visual complexity.[2]

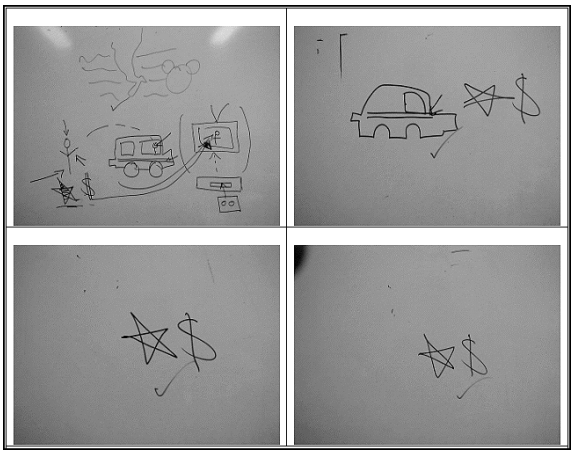

Figure 3: A Sequence of 'Robert deNiro's from the Concept Drawing Task

Experiments to evaluate the effects of these contractions on the intelligibility of the drawings for non-participants are currently in progress. However, it appears likely that they will have a substantial effect. The first drawing in Figure 3 has a number of elements that might allow a non-participating observer to guess the identity of the individual depicted. For example it includes a sketch map of Italy, sketches of a TV and VCR, and an image of a taxi (which refers to a de Niro film). By contrast, the last sketch in the sequence, consisting of a star and a dollar sign would be much harder to decipher.

Effects of Direct Interaction. Experiments with the Music Drawing task have investigated the influence of level of communicative interaction between participants on the type of drawing (Abstract or Figurative) that they produce. The basic contrast is between an interactive and non-interactive version of the task (Healey et al., 2001; Healey, Swoboda, Umata and Katagiri, forthcoming). In the

[2]The analysis of visual complexity is based on a psychophysical measure developed by Pelli Burns Farell and Moore (in press) and is based on the formula: Complexity = Perimeter^2 / Ink.

non-interactive version, subjects alternate between acting as Drawer and Chooser on each trial and the whiteboard is configured to prevent the Chooser from drawing. In this version of the task each trial approximates to a single turn in the communicative exchange. In the interactive version the task is altered so that both members of a pair draw at the same time. They have one piece of music each and must determine, using only drawing, whether their pieces are the same or different. In this case there is a richer communicative exchange. In addition to producing drawings of their pieces, subjects employ devices such as arrows, underlining, and circling to query and revise various aspects of their drawings. Each trial in the interactive task thus approximates to a number of 'conversational' turns.

The effect of the difference in level of communicative interaction can be seen in Table 1 (the 'Composite' category refers to drawings that combine Figurative and Abstract elements). Where both members of a pair can interact directly on the whiteboard, they rely primarily on the Abstract drawings. In the non-interactive task, where they are alternating between drawing and choosing, they rely primarily on Figurative drawings.

Table 1: Distribution of Drawing Types in the Music Drawing Task

	Drawing Type		
Task	Abstract	Figurative	Composite
Interactive	59%	21%	16%
Non-Interactive	27%	64%	8%

Further evidence for the importance of direct interaction comes from analysis of the logs of drawing activity (Healey, Swoboda, Umata and Katagiri, forthcoming). The Abstract and Figurative drawing types are not distinguishable in terms of the number of lines or ink (pixels) involved in producing them, nor in terms of the accuracy of responses associated with their use. Considerations of the efficiency or effectiveness of the two drawing types alone do not appear to explain their pattern of use. However, drawing activities overlap approximately 20% more when subjects produce Abstract drawings than when they produce Figurative drawings. This suggests it is the availability of specific mechanisms of communicative interaction, such as the circling and underlining of each others drawings, that is critical to the co-ordinated use of the Abstract drawings.

Community-based Conventions. Perhaps the most interesting parallel between graphical and verbal dialogue comes from experiments on the emergence of graphical conventions in experimental 'sub-communities' (cf. Garrod and Doherty, 1994).

Data from an unpublished experiment with the Music Drawing task demonstrates that, for this task at least, the patterns of co-ordination in drawing style that emerge within sub-communities are specific to those sub-communities (cf. Healey, 1997). The experiment takes place in two phases. In the first 'convergence' phase experimental sub-communities consisting of sub-groups of six people are formed. Subjects themselves are unaware of this sub-group manipulation, from their perspective the experiment consists of a series of rounds of Music Drawing with a different partner each time. During the convergence phase, the composition of pairs is controlled so that they are always made up of individuals from within the same sub-group. This continues for four rounds thus allowing for a history of interactions to build up within each sub-group. On each round subjects perform the interactive version of the Music Drawing Task for 12 trials.

The second, experimental, phase occurs in the fifth round. In this round two conditions are compared; same-group pairs who are composed, as before, of subjects from within a single sub-group and cross-group pairs who are composed of subjects drawn from different subgroups.[3] Same-group and cross-group pairs have equivalent task experience and, as noted, are unaware of any sub-group manipulation. Nonetheless they are reliably different in their use of the Drawing types. Multinomial regression analysis shows a reliable effect of the group manipulation on the distribution of Drawing types ($Chi^2_{(3)}$=25.44, p=0.00, n=516). The percentages are shown in Table 2.

Table 2: Use of Drawing Types in Pairs Drawn from the Same or Different Subgroups

	Drawing Type		
Task Version	Abstract	Figurative	Composite
Same-group	62.7%	11.1%	18.1%
Cross-group	41.3%	32.9%	15.8%

These results indicate that the co-ordination on particular drawing types that develops within the experimental sub-communities is community-specific. Subjects in the cross-group interactions use a more mixed profile of drawing types. This suggests that the graphical conventions established within sub-communities do not readily transfer to interaction outside those communities. Healey (1997) reports parallel results for verbal dialogues about spatial locations. In this case the types of spatial referring expressions established within sub-communities

[3]The original design employed three experimental subgroups but for ease of exposition only two are reported here.

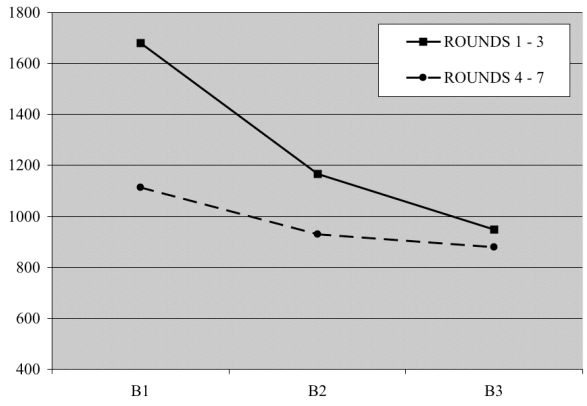

Figure 4: Reduction in Visual Complexity of Concept Drawings with Repetition in a Round (B1-B3 = Blocks of recurring items within a round)

during the corresponding convergence phase also proved unstable in 'cross-group' interactions.

Data from a community-based version of the Concept Drawing Task also suggests parallels between communities of graphical and linguistic communicators. The task requires a community group of 8 participants to communicate with each of the other 7 over an extended period of time. In the first round of the experiment they work in 4 pairs with both participants drawing each concept 3 times over the course of the round. In the second round the 8 participants are re-paired and again draw the concepts 3 times. After each round they are re-paired again until every participant has encountered each of the others once and only once.

Figure 4 shows how drawings become increasingly simple (on the Pelli et al. measure) as the experiment proceeds. In the first 3 rounds this simplification process occurs across repetitions of the drawings (shown along the x axis of the figure). However, as the shared interaction within the community begins to develop (i.e., after round 4) the initial drawings in a round become as simple as the final drawings in the round. A similar pattern of results emerges for the communicators accuracy at identifying the concepts conveyed by their partners drawings. These findings are consistent with the idea that as a community becomes established through a common history of interaction so the drawings become conventionalised within the community: Drawings become simpler and more readily interpreted by the members of the community.

The implication of these results is that the processes which establish the conventions for producing and interpreting drawings and verbal descriptions operate in a manner that is directly tied to the character and pattern of interactions in which they were developed and used.

Discussion

The aim of providing an overview of a number of experimental results means that much important detail has been elided from the descriptions of experiments and results provided above. Nonetheless, the results summarised above consistently point to the importance of interactional context in graphical communication.

Like referring expressions in conversation, the form and interpretation of drawings is systematically influenced by the character of the interaction in which they occur. Participants in interactions show a strong tendency to match each others representational style and type. If items recur in an interaction, pairs also tend to develop increasingly abbreviated ways of representing them that are difficult for third parties to interpret. These patterns of change in the form of drawings obtain independently of the particular concept or item being represented. In addition to these basic co-ordination processes of entrainment and abbreviation, there is evidence that level of direct graphical interaction available to participants affects the form of representations they use. In particular, the ability to mark up and modify elements of each others' drawings appears to be important to the sustained use of more abstract representations. Lastly, this paper has presented evidence that interactions within sub-communities lead to the development of community-specific conventions for graphical interaction.

The programmatic rationale for investigating tasks, such as those described above, that involve exclusively graphical communication is the potential they offer for comparison with other modes of interaction. The results summarised above suggest significant parallels between the mechanisms that underpin communicative co-ordination in exclusively graphical and verbal exchanges. As noted above, some of these findings can be accounted for in terms of the collaborative model of grounding (Clark and Wilkes-Gibbs 1986; Clark, 1996) and input-output coordination model (Garrod and Anderson 1987, Garrod and Doherty 1994). The importance of interactional mechanisms, such as localisation, to graphical communication also suggests possible parallels with the mechanisms of conversational repair (Sacks, Schegloff and Jefferson 1974; Schegloff 1992). The viability of applying these explanations to the details of graphical communication is the subject of further work.

Acknowledgements

We gratefully acknowledge the support of ESRC/EPSRC under the PACCIT programme for the project MAGIC: Multimodality and Graphics in Interactive Communication (L328253003), and ATR Media Information Science Laboratories. We are especially grateful to James King, Nik

Swoboda, Ichiro Umata and Yasuhiro Katagiri for their contributions to this research. An earlier version of this paper was presented under the title "Interactional Context in Sketch Understanding" at the AAAI Spring Symposium, Stanford, 25th-27th of March, 2002.

References

Brannigan, H., Pickering, M., & Cleland, A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(B), 13-25.

Clark, H. H. (1996). *Using language.* Cambridge: Cambridge University Press.

Clark, H. H. (1998). Communal lexicons. In K. Malmkjoer & J. Williams (Eds.), *Context in language learning and language understanding* (3rd ed., pp. 63–87). Cambridge: CUP.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Engle, R. (1998). Not channels but composite signals: Speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. In M. Gernsbacher & S. Derry (Eds.), *Proceedings of the 20th annual conference of the cognitive science society* (pp. 321–326).

Garrod, S. C., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*, 181–218.

Garrod, S. C., & Doherty, G. (1994). Conversation, coordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*, 181–215.

Healey, P. (1997). Expertise or expert*ese*?: The emergence of task-oriented sub-languages. In M. Shafto & P. Langley (Eds.), *Proceedings of the 19th annual conference of the cognitive science society* (pp. 301–306).

Healey, P., & Peters, C. (2001, 18th-20th April). *Notes on turn-taking and topic in drawing-in-interaction.* (Paper presented at the *The 4th International Workshop on Gesture and Sign Language based Human-Computer Interaction* City University, London)

Healey, P., Swoboda, N., & King, J. (forthcoming). *A tool for performing and analysing experiments on graphical communication.* (Paper to be presented at *HCI2002: The 16th British HCI Group Annual Conference*, September 2nd- 6th, South Bank University, London)

Healey, P., Swoboda, N., Umata, I., & Katagiri, Y. (2001). Representational form and communicative use. In J. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the cognitive science society* (pp. 411–416).

Healey, P., Swoboda, N., Umata, I., & Katagiri, Y. (forthcoming, October). *Graphical representation in graphical dialogue.* (Paper to appear in a special issue of the *International Journal of Human Computer Studies* on Interactive Graphical Communication)

Hymes, D. (1972). Models of the interaction of language and social life. In J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 35–71). New York: Holt, Rinehart and Wilson.

Krauss, R., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In A. Higgins, E.T. nad Kruglanski (Ed.), *Social psychology: Handbook of basic principles* (pp. 655–701). London: Guildford Press.

Neilson, I., & Lee, J. (1994). Conversations with graphics: implications for the design of natural language/graphics interfaces. *International Journal of Human-Computer Studies*, *40*, 509–541.

Pelli, D. G., Burns, C. W., Farell, B., & Moore, D. C. (in press). Identifying letters. *Vision Research.*

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation. *Language*, *50*, 696–735.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*(2), 185–213.

Schegloff, E. A. (1992). Repair after the next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, *97*(5), 1295–1345.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*, 211–232.

Sommers, P. van. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes.* Cambridge: Cambridge University Press.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, *31*, 183–194.