The fitness and functionality of culturally evolved communication systems

Nicolas Fay[1,*], Simon Garrod[2] and Leo Roberts[1]

[1]*The University of Western Australia, Crawley, WA 6009, Australia*
[2]*The University of Glasgow, Glasgow G12 8QQ, UK*

This paper assesses whether human communication systems undergo the same progressive adaptation seen in animal communication systems and concrete artefacts. Four experiments compared the fitness of ad hoc sign systems created under different conditions when participants play a graphical communication task. Experiment 1 demonstrated that when participants are organized into interacting communities, a series of signs evolve that enhance individual learning and promote efficient decoding. No such benefits are found for signs that result from the local interactions of isolated pairs of interlocutors. Experiments 2 and 3 showed that the decoding benefits associated with community evolved signs cannot be attributed to superior sign encoding or detection. Experiment 4 revealed that naive overseers were better able to identify the meaning of community evolved signs when compared with isolated pair developed signs. Hence, the decoding benefits for community evolved signs arise from their greater residual iconicity. We argue that community evolved sign systems undergo a process of communicative selection and adaptation that promotes optimized sign systems. This results from the interplay between sign diversity and a global alignment constraint; pairwise interaction introduces a range of competing signs and the need to globally align on a single sign-meaning mapping for each referent applies selection pressure.

Keywords: graphics; communication; signs; cultural evolution; fitness

1. INTRODUCTION

Like everything else in the natural world, communication systems evolve: their signals adapt to best fit the circumstances of the communication. For example, the 'whine-plus-chuck' mating call of the male Panamanian frog *Physalaemus pustulosus* is perfectly adapted to the communicative situations in which it is used. On one hand, its pitch lies within the range of the best-hearing frequencies of the female whom it attracts; on the other hand, the female attracting 'chuck' part of the call is just short enough to make it difficult for predatory bats to localize (Seyfarth & Cheney 2003). This kind of biological (i.e. genetic) adaptation may also be seen in the evolution of human languages. For example, Dediu & Ladd (2007) showed that the marked geographical distribution of tone languages is a result of the distribution of recently evolved alleles of the brain growth and development genes *ASPM* and *Microcephalin*. They suggested that this reflects the influences of the genes on the ability of their owners to easily acquire the tone languages.

However, human communication systems might also evolve through processes of cultural evolution (Christiansen & Kirby 2003; Kirby *et al.* 2007). Whereas biological adaptation optimizes the language learning machinery via innate learning biases (Pinker 1994;

Dediu & Ladd 2007), cultural transmission, the historical transmission of languages across generations of learners, could optimize language via linguistic selection (see Kirby & Hurford (1997) for a computer simulation).

In this paper, we explore the analogy between biological and cultural evolution by testing whether cultural transmission tends to produce an optimized, or fit, system of signs in the way that biological evolution shapes the mating call of the Panamanian frog. In so doing, we assess the veracity of the functionalist view of language, which argues that language has evolved to support precise and efficient communication (Pinker & Jackendoff 2005).

Historical and laboratory studies of cumulative cultural evolution, the process by which knowledge accumulates across generations, support the functionalist perspective. Later generations improve upon the solutions provided by earlier generations, eventually arriving at solutions that no single individual could produce on their own. According to Tomasello (1999; Tomasello *et al.* 2005), this incremental improvement in the quality of solutions (ratcheting up) is based on social learning mechanisms unique to humans. Technological evolution offers a compelling historical example of cumulative cultural evolution. Employing an organic–mechanical analogy, Basalla (1988) proposed that human artefacts undergo a Darwinian process of survival of the fittest. Specifically, artefacts that are best suited for certain tasks survive, and are subject to gradual modifications that improve their functionality. This is seen in the progressive improvement of the hammer, evolving from a crudely shaped

pounding stone to today's claw and ball-pein hammers. A similar outcome is evident under controlled laboratory conditions, where performance in producing artefacts improves, or ratchets up, over successive generations (e.g. distance travelled by paper aeroplanes and height of spaghetti tower constructions; Caldwell & Millen 2008; see also Caldwell & Millen 2008).

It is an open question whether these evolutionary principles apply to the development of symbolic artefacts (e.g. linguistic and other sign systems). In particular, it is unclear whether linguistic systems undergo the progressive improvement evident in biological systems and concrete artefacts. As Kirby (2002, p. 194) observed, 'We cannot take it for granted that either learning or cultural evolution are adaptive mechanisms that seek optimal solutions with regard to communication, however intuitively appealing that may appear'. This view was echoed by Plotkin (2002), who noted that 'social constructions' are a product of shared agreement, and as such may require a fundamentally different explanation from concrete traits such as technological artefacts (cited in Mesoudi *et al.* 2006).

2. ASSESSING THE EVOLUTION OF HUMAN COMMUNICATION SYSTEMS

Despite the recent resurgence of interest in how present-day languages have evolved (e.g. Lieberman *et al.* 2007; Pagel *et al.* 2007), little is known about what drives this evolution. In particular, it is still unclear whether languages increasingly adapt to fit the needs of speakers and listeners. It is difficult to assess the fitness of evolved linguistic systems because there is little available empirical data; almost all of the systems used today originated in the prehistoric past. However, there has been some interesting research on recently evolved sign languages (e.g. Kegl *et al.* 1999; Goldin-Meadow 2003; Sandler *et al.* 2005). One way of overcoming the lack of linguistic fossils for spoken languages is to use computer simulations of communicating agents to test different hypotheses about how language might evolve (Kirby & Hurford 2002; Steels *et al.* 2002; Barr 2004). While computational models have identified several parameters important for language evolution, they do not speak to the fitness of the evolved linguistic symbols. Steels *et al.* (2002) showed that a stable lexicon is established via the interactions of a community of computer agents, but their results do not explain, for example, why 'wogglesplat' was selected among other signs to convey a particular meaning (e.g. red triangle).

An alternative approach has been to study how present-day humans perform novel communication tasks without access to a previously established sign system. Most of the tasks that have been used to do this involve graphical communication with participants having access to various drawing media (Galantucci 2005; Garrod *et al.* 2007; Healy *et al.* 2007). The rationale for using graphical communication tasks to study the emergence and evolution of sign systems is that, with the exception of writing and reading, graphical communication is extremely rare. So graphical communication tasks allow us to study how people adapt to new communication media and how graphical sign systems emerge and evolve over time.
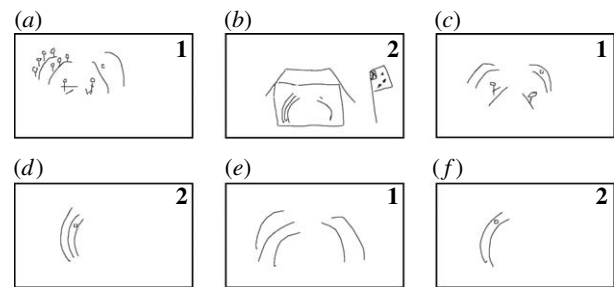

Figure 1. Drawing refinement and alignment for the concept 'Parliament' across six games between a pair of interlocutors playing the Pictionary task (adapted from Fay *et al.* in preparation): (*a*) game 1, (*b*) game 2, (*c*) game 3, (*d*) game 4, (*e*) game 5, (*f*) game 6. Participant numbers are given in bold on the top right of the drawing.

A consistent finding across studies is the crucial role of feedback and interaction to the development of graphical sign systems. In Garrod *et al.* (2007), participants communicated a series of predetermined concepts by drawing on a standard whiteboard. Like the game Pictionary, participants were not allowed to speak or use text in their drawings, forcing them to create a novel sign system. In one condition, pairs of participants graphically communicated a set of recurring concepts (e.g. Art gallery, Drama, Arnold Schwarzenegger and Television), alternating between drawing and identifying roles from game to game. Figure 1 illustrates the changing form of the sign representing 'Parliament' across six games.

What begins as an iconic depiction of Parliament, a figurative illustration of the debating chambers, develops, through a process of local adaptation and entrainment, into a simplified symbolic form (two lines plus a circle). Not only are participants' drawings refined across games, but they also become increasingly similar, or aligned. Crucially, simple repetition of drawings was insufficient to produce simplification and abstraction. This occurred only when there was a feedback from the addressee. Garrod *et al.* (2007) argued that interactive graphical communication allows participants to develop shared symbolic representations from what started out as primarily iconic representations through a 'grounding' process similar to that found in interactive spoken communication (Clark 1996). Further research using the Pictionary task contrasts this local process with the global evolution of a 'visual lexicon' within a community of interlocutors (Fay *et al.* in preparation).

Four 8-person laboratory communities, or microsocieties, were created via the one-to-one interactions of partners drawn from the same pool. Participants played six consecutive games with a partner, where each game contained the same to-be-communicated items (16 targets plus four distracters, presented in a different random order on each game) that were known to both partners. As in the previous example, drawing and identifying roles alternated from game to game. Participants then switched partners and played a further six games with a new partner, and continued to do so until they had interacted with each of the other community members (table 1 displays the sequence of partner interactions in each community). Communities were

Table 1. The sequence of partner interactions in each community. (Within each 'round', participants played six consecutive games of the Pictionary task with a different partner.)

round	pair composition			
1	1 and 2	3 and 4	5 and 6	7 and 8
2	1 and 4	3 and 2	5 and 8	7 and 6
3	1 and 6	3 and 8	5 and 2	7 and 4
4	1 and 8	3 and 6	5 and 4	7 and 2
5	1 and 3	2 and 4	5 and 7	6 and 8
6	1 and 5	2 and 6	3 and 7	4 and 8
7	1 and 7	2 and 8	3 and 5	4 and 6

designed such that a conventional communication system could be established by the time participants encountered their fourth partner. For instance, assume person 2 adopts person 1's sign system (round 1, table 1), and that person 2 then influences person 3 (round 2). If person 8 aligns with person 3 (round 3), persons 1 and 8 will share a similar communication system (round 4) despite having never directly interacted.

The community condition was contrasted with an isolated pair condition, in which participants interacted with the same partner over the same number of games (i.e. 42 games; see Garrod & Doherty (1994) for a natural language analogue). The task was administered using a virtual whiteboard tool (Healy *et al.* 2002), with each participant seated at a computer terminal and drawing input and item selection made via a standard mouse. Crucially, participants were unaware of the identity of their partner in any round. Figure 2 illustrates the global and local evolution of the sign representing 'Brad Pitt' within a single community and a corresponding number of isolated pairs.

The first drawings of Brad Pitt (figure 2, round 1) illustrate the diversity of graphical signs; some indicate his American origins, others his frequent casting as a ladies man, while others use the rebus principle to represent part of the test item (community members 5 and 6 draw a large hole in the ground to convey a 'pit', whereas isolated pair member 4 draws an arrow pointing at an arm pit). Drawing diversity at round 1 in the community condition contrasts sharply with drawing uniformity at round 7, where all community members have globally converged on a refined version of person 5's initial pit drawing. Unlike community members, isolated pairs locally converged on a shared sign system, but globally diverged across games. Note that in both conditions groups arrived at a series of signs of equal visual complexity (see Fay *et al.* in preparation).

In this paper, we investigate whether the signs (i.e. drawings) produced by communities are better adapted for use in the larger population than those produced by isolated pairs. If these signs undergo systematic adaptation, then we would expect signs evolved by isolated pairs to be fit only for that pair, whereas community evolved signs should be fit for the larger population from whom the community had been created. We test this hypothesis by contrasting the accuracy and ease with which community and isolated pair evolved signs can be learned (experiments 1 and 2) and detected (experiment 3) by new subjects drawn

from the same population. Experiments 1–3 can be thought of as fractionating 'levels of meaningfulness' of the signs: decoding of learned signs assesses the strength of the stored sign-meaning mapping (high meaningfulness); encoding of learned signs addresses the effort required to discriminate between the signs themselves (low meaningfulness); and detection determines the ease with which the signs are perceived (no meaning).

Clearly an effective sign should be easily detected, efficiently encoded into memory and its meaning should be accurately and efficiently derived from the sign. However, it may be that community evolved signs are superior to those that develop among isolated pairs because the signs preserve more salient, concrete information, despite their comparable visual complexity. In other words, they have more residual iconicity. If greater iconicity distinguishes community from isolated pair evolved signs (i.e. the strength of the sign-meaning mapping), then the benefits of community signs should be most clearly seen in experiment 1 (decoding). Experiment 4 provides a direct test of sign iconicity by testing naive overseers' ability to guess the meaning associated with community and isolated pair evolved signs. Thus, the current study tests the intuitive hypothesis that the product of cultural evolution, or glossogeny, is an optimized sign system, a position consistent with the functionalist perspective of language.

3. LEARNABILITY STUDIES

The first two experiments investigate the learnability of signs evolved by communities as compared to those developed by isolated pairs. Experiment 1 assesses the relative decoding of the two kinds of meaningful sign. Experiment 2 assesses their learned discriminability.

(a) *Experiment 1. Decoding community and isolated pair evolved signs*
(i) *Methods*
Participants and apparatus
Thirty-two undergraduate psychology students participated in exchange for payment. Participants were tested individually in sessions lasting 45 min. Stimuli were presented and controlled by a personal computer with a monitor refresh rate of 100 Hz (the same apparatus was used in experiments 2–4).

Materials and design
Stimulus materials were drawn from Fay *et al.* (in preparation). The stimuli consisted of 512 community drawings (32 participants randomly allocated to 4×8-person communities) and 512 isolated pair drawings (32 participants randomly allocated to 16 isolated pairs). Half the drawings were sampled at game 1 of round 1 (i.e. pre-interaction; 16 concepts×32 pairs) and the other half at game 1 of round 7 (i.e. post-interaction; 16 concepts×32 pairs). Images were sampled such that each participant was presented with images produced by a community pair and an isolated pair at round 1 or round 7. This equated to 32 images per participant (16 community produced images at round 1/7 and 16 isolated pair generated images at round 1/7). Thus, the corpus was sampled once across participants. A mixed design was used.

Figure 2. Drawing refinement and alignment for the concept 'Brad Pitt' among (*a,b*) a community of interlocutors and (*c,d*) between isolated pairs at round 1 (*a,c*) and round 7 (*b,d*) (game 1) of the Pictionary task (adapted from Fay *et al.* in preparation). Participant numbers are given in bold on the top right of the drawing.

Procedure

Participants were tested individually in a quiet testing booth. The experiment began with a training phase, where the participants learned the identity of each image (32) before progressing to a two alternative forced-choice reaction time task. During training, participants viewed each image and its associated label (e.g. 'Microwave' and 'Soap Opera'; presented in a random order), pressing the space bar to progress to the next image–label pair. At test, each image was presented and participants were cued to select the associated label from an adjacent list. Identification accuracy of 80 per cent or better was required for participants to proceed to the reaction time task. If required, the training phase was repeated until the participant achieved 80 per cent accuracy or better on test.

The timed decoding task required participants to decide whether a presented label matched a previously presented image. Each trial began with the presentation of a fixation cross in the centre of the screen (500 ms). Next the target image was presented (40 ms), followed immediately (0 ms onset) by a mask (50 ms). Each mask was a scrambled version of the most complex target image (measured in pixels). A matching or mismatching label was then presented in the centre of the screen, to which participants responded either match (by pressing the 'f' key) or mismatch (by pressing the 'j' key). A feedback beep indicated whether their response was correct or incorrect (see figure 3 for a sample trial sequence).

Half the trials were 'match' trials, where the target image agreed with the associated label. The remaining trials were 'mismatch' trials (e.g. a drawing of
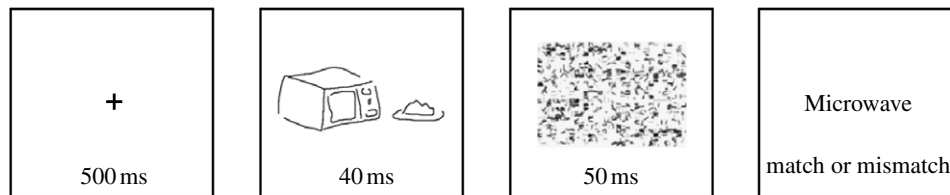
Figure 3. A sample trial sequence from experiment 1.

Microwave followed by the label Parliament). Each image was displayed eight times: four times in a match trial and four times in a mismatch trial. Images and their associated labels were presented in a random order throughout. Participants completed 256 trials (16 concepts×2 pairs×2 trial types×4 repetitions). The computer recorded their accuracy and response latency for each trial.

Results

All participants successfully completed the training phase before progressing to the reaction time task. A majority learned the image–label pairing to the pre-determined criterion (80% accuracy) at the first attempt (21, or 66%), with a few requiring a second attempt (8, or 25%) and fewer still requiring a third attempt (3, or 9%). Round 1 and round 7 community and isolated pair images were equally well learned at training (88% accuracy).

Figure 4 displays the mean hit rate (% correct responses) and response latencies (in milliseconds) for community and isolated pair evolved signs at round 1 and round 7 (image–label matching trials). Performance at round 1 is equivalent. At round 7, community evolved signs are more accurately recognized and more efficiently processed than isolated pair developed signs. ANOVA confirms these observations.

Participants' mean hit rates were entered into a mixed-design ANOVA, treating group (community and isolated pair) as a within-subjects factor and round (1 and 7) as a between-subjects factor (for all F and t values reported $p < 0.05$ unless otherwise stated). This returned a main effect of group ($F_{1,30} = 7.28$, $\eta_p^2 = 0.20$), but no effect of round ($F_{1,30} = 2.76$, $p > 0.05$, $\eta_p^2 = 0.08$). However, the main effect of group was mediated by a reliable group by round interaction ($F_{1,30} = 4.12$, $\eta_p^2 = 0.12$). This was due to the simple effect of group at round 7 ($F_{1,30} = 11.17$, $d = 1.08$), with no such effect at round 1 ($F < 1$). Identical findings were returned when the ANOVA was repeated using participants' d scores.

The same ANOVA was carried out on participants' response latencies. Mean response times were calculated after the removal of times 2.5 standard deviations from the condition median. These extreme scores were replaced by values corresponding to the median plus or minus 2.5 standard deviations. This accounted for 2.7 per cent of the data. ANOVA returned a main effect of group ($F_{1,30} = 12.46$, $\eta_p^2 = 0.29$), but no effect of round ($F < 1$). Again, the main effect of group was mediated by a reliable group by round interaction ($F_{1,30} = 9.85$, $\eta_p^2 = 0.25$). The interaction reflects the simple effect of group at round 7 ($F_{1,30} = 22.23$, $d = 0.47$), with no difference at round 1 ($F < 1$).
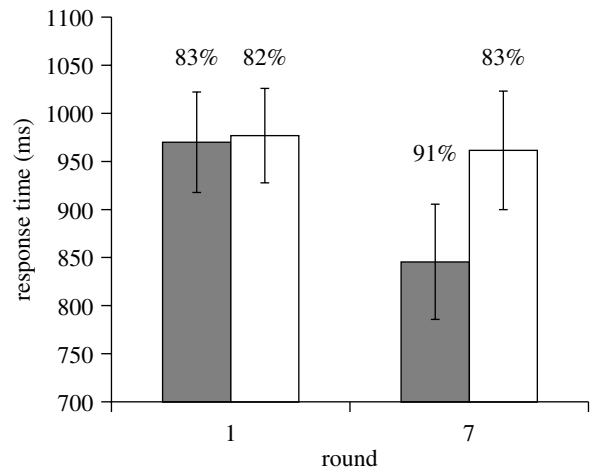
Figure 4. Mean hit rate (%) and response latency (ms) for community (grey bars) and isolated pair (white bars) evolved signs in image–label matching trials at round 1 and round 7. Error bars indicate the standard error of the mean.

In summary, community evolved signs (round 7) offer a substantial learning advantage over signs that locally develop among isolated pairs. Furthermore, these signs are more rapidly accessed when compared with isolated pair developed signs of equal visual complexity. To establish the validity of the iconicity-based account (i.e. that the accessibility benefits seen for community evolved signs are a function of their greater residual iconicity), two competing explanations must be ruled out: encoding and detection. Ease of sign encoding and detection are tested in experiments 2 and 3.

(b) Experiment 2. Encoding community and isolated pair evolved signs

The decoding benefits associated with community evolved signs (experiment 1) may be attributable to more efficient sign encoding (i.e. community evolved signs are more efficiently encoded in working memory) and greater sign discriminability (i.e. community evolved signs are more distinct and are therefore less confusable than isolated pair developed signs). Experiment 2 tested this possibility by comparing the ease with which community and isolated pair evolved signs are encoded and distinguished from one another.

(i) Method
Participants
Thirty-two undergraduate psychology students participated in exchange for payment. Participants were tested individually in sessions lasting 45 min.

Materials and design
Stimulus materials were again drawn from Fay *et al.* (in preparation). Eight images were sampled from each community pair and isolated pair at round 1 and round 7. Images were sampled such that each participant

was exposed to all 16 concepts (four targets and four distracters sampled from a community pair and an isolated pair at round 1 or round 7), with no duplication of item types. This meant that half the corpus was used. A mixed design was employed.

Procedure

The experiment consisted of a training phase in which participants learned eight target images followed by an inspection-time task. At training, participants viewed eight target images presented in a random sequence, pressing the space bar to proceed from one target to the next. A recognition memory test followed where participants were cued to identify each image as a target or distracter (eight targets plus eight distracters). Memory performance exceeding 80 per cent allowed progression to the inspection-time task.

The inspection-time task required participants to recognize images (as targets or distracters) presented at varying exposure durations. A parameter estimation by sequential testing algorithm (Taylor & Creelman 1967) determined the minimum exposure duration required by each participant to respond at 70 per cent accuracy (see Treutwein (1995) for a review). Individual staircases lasting 50 trials were performed on each image. Each trial consisted of the presentation of an image for the determined inspection time, followed by a mask (0 ms onset) that remained on the screen until participants identified the image as a target (by pressing the 'f' key) or a distracter (by pressing the 'j' key). The mask was identical to that used in experiment 1. A feedback beep informed participants if their response was correct or incorrect. Participants completed 800 trials in total (16 images × 50 trials). Image presentation was randomized throughout.

Results

All participants successfully completed the training phase at the first attempt. Community and isolated pair target images were equally well learned at round 1 and round 7 (97% accuracy). Participants' mean inspection times were calculated across the last five trials for each target image (community or isolated pair at round 1 or round 7). Inspection times were computed after the removal of times 2.5 standard deviations from the condition median. These extreme scores were replaced by values corresponding to the median plus or minus 2.5 standard deviations. This accounted for 3.4 per cent of the data.

Figure 5 displays the mean inspection time (in milliseconds) required to identify each target image to the predetermined criterion (70% accuracy). There was no difference between conditions. This was confirmed by ANOVA (same design as experiment 1). There was no effect of group ($F_{1,30} = 2.78$, $p > 0.05$, $\eta_p^2 = 0.08$), round ($F < 1$) or group by round interaction ($F < 1$). The similar inspection times for round 1 and round 7 signs indicate that the greater visual complexity of round 1 signs did not slow sign encoding, suggesting that participants needed only to encode part of the round 1 signs for successful recognition. More importantly, the near-identical encoding efficiency of round 7 community and isolated pair images indicates that the decoding benefits of community evolved signs
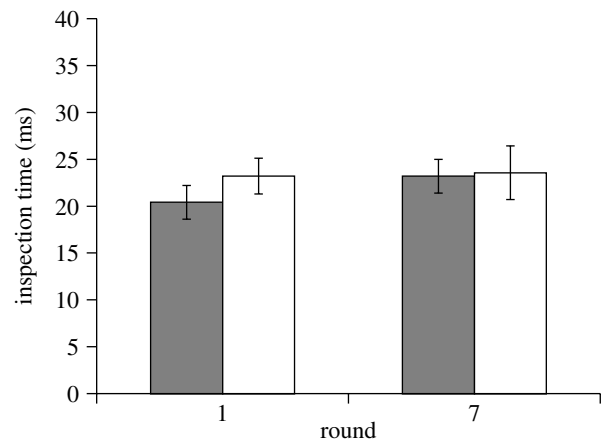

Figure 5. Mean inspection time (ms) required to correctly recognize round 1 and round 7 community (grey bars) and isolated pair (white bars) target images on 70 per cent of the trials. Error bars indicate the standard error of the mean.

(experiment 1) cannot be attributed to more efficient sign encoding or discrimination.

(c) *Experiment 3. Detecting community and isolated pair evolved signs*

If community evolved signs are easier to detect than isolated pair developed signs, then this would provide an alternative explanation of the decoding benefits seen for community evolved signs (experiment 1). This possibility was tested in experiment 3.

(i) *Method*

Participants

Sixteen undergraduate psychology students participated in exchange for payment. Participants were tested individually in sessions lasting 1 hour.

Materials and design

Stimulus materials were identical to those used in experiment 1. Images were sampled such that each participant was presented with images generated by two community pairs (from different communities) and two isolated pairs, randomly sampled at round 1 and round 7. This equated to 128 images per participant (16 concepts × 2 community pairs × 2 isolated pairs × 2 rounds). Thus, the corpus was sampled twice across participants. A within-subjects design was used.

Procedure

Participants completed a two alternative forced-choice task in a quiet testing booth. Each image was presented once at each of three exposure durations (10, 20 or 30 ms), with a corresponding number of target-absent trials (i.e. 128 target-present trials plus 128 target-absent trials × 3 exposure durations). Images were presented in a random order. Each trial began with the presentation of a fixation cross in the centre of the screen (500 ms). Next a target image or blank screen was presented (10, 20 or 30 ms), followed immediately (0 ms onset) by a mask (50 ms). The mask was identical to that used in experiments 1 and 2. Participants were cued to respond whether the target was present (by pressing the 'f' key) or absent (by pressing the 'j' key).
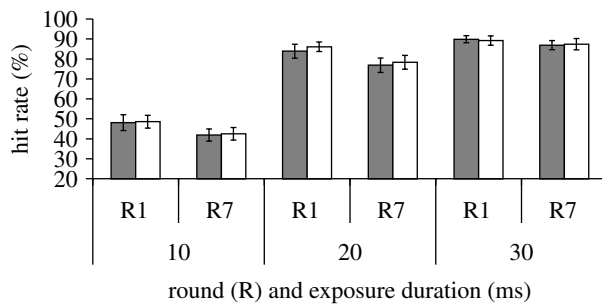
Figure 6. Mean hit rate (%) for community (grey bars) and isolated pair (white bars) signs at round 1 (R1) and round 7 (R7) at 10, 20 and 30 ms exposure durations. Error bars indicate the standard error of the mean.

Results
Figure 6 displays the mean hit rate (% correct responses in target-present trials) for each condition at 10, 20 and 30 ms exposure durations. Community and isolated pair evolved signs were detected equally well at each exposure duration, although in both conditions round 1 images were more accurately detected than round 7 images.

Participants' mean performance scores (% correct) were entered into an ANOVA that treated group (community and isolated pair), round (1 and 7) and exposure duration (10, 20 and 30 ms) as within-subject factors. This returned a main effect of round ($F_{1,15}=10.99$, $\eta_p^2=0.42$) and exposure duration ($F_{1,15}=129.31$, $\eta_p^2=0.90$), but no effect of group ($F<1$). There were no interaction effects ($F_s<2.82$). Further analysis of the main effect of exposure duration confirmed a large improvement in detection rates between 10 and 20 ms ($t_{15}=12.54$, $d=3.23$) and a smaller improvement between 20 and 30 ms ($t_{15}=2.45$, $d=0.71$).

Thus, participants were better able to detect the presence of an image at longer exposure durations. Not surprisingly, participants were also more successful at detecting the more visually complex round 1 images. Identical findings were returned when the ANOVA was repeated using participants' d scores. The equivalent detection rates for community and isolated pair images confirm that the decoding benefits seen for community evolved signs (experiment 1) cannot be attributed to ease of detection.

(d) Experiment 4. The transparency of community and isolated pair evolved signs
Experiment 4 provides a direct test of the hypothesis that community evolved signs have greater residual iconicity than those developed by isolated pairs. To test this, we assessed the degree to which naive observers could guess the original meaning of the two kinds of signs.

(i) Method
Participants and apparatus
Thirty-two undergraduate psychology students participated in exchange for payment. Participants were tested individually in sessions lasting 30 min.

Materials and design
Stimulus materials were again drawn from Fay *et al.* (in preparation). Unlike experiments 1–3, where participants were shown static images, in experiment 4
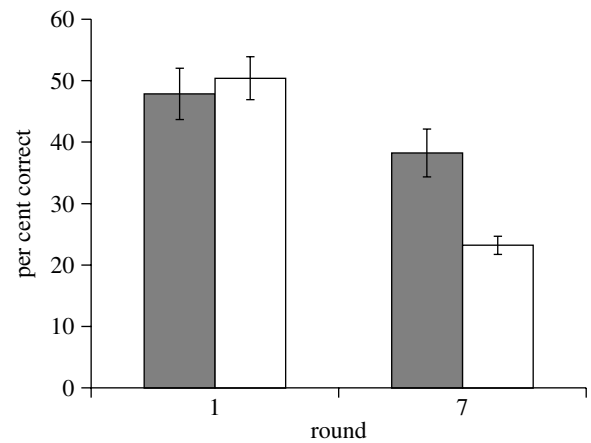
Figure 7. Overseers' mean identification accuracy (%) for community (grey bars) and isolated pair (white bars) signs at round 1 and round 7. Error bars indicate the standard error of the mean.

images were animated, replicating the dynamic drawing construction experienced by the actual matcher. The virtual whiteboard tool (Healy *et al.* 2002) used by Fay *et al.* (in preparation) enables pixel-by-pixel playback of the drawing activity within each experimental trial. Trial playback was converted to QuickTime animations and used as stimuli in experiment 4. Each participant attempted to identify the meaning of the animated drawings produced by a community pair and an isolated pair at round 1 or round 7 (32 animations in total). The corpus was sampled once across participants. A mixed design was used.

Procedure
Participants completed the task individually in a quiet testing booth. Trials were initiated with the presentation of a fixation cross (500 ms) followed by animation playback. Participants then tried to identify the referent of the animated drawing (by key press) from an adjacent list of 20 concepts (e.g. 'Homesick', 'Cartoon' and 'Computer Monitor'). The animated drawings were presented in a random order.

Results
Figure 7 displays overseers' mean identification rate (%) for community and isolated pair evolved signs at round 1 and round 7. Identification accuracy at round 1 is equivalent, whereas at round 7 community evolved signs are more accurately identified. This was confirmed by ANOVA.

Participants' mean identification accuracy scores were entered into a mixed-design ANOVA as per experiments 1 and 2. This returned a main effect of group ($F_{1,30}=5.06$, $\eta_p^2=0.14$), round ($F_{1,30}=21.35$, $\eta_p^2=0.42$) and a reliable group by round interaction ($F_{1,30}=10.05$, $\eta_p^2=0.25$). The interaction is explained by the simple effect of group at round 7 ($F_{1,30}=14.68$, $d=1.27$), with no such effect at round 1 ($F<1$). Thus, community evolved signs retain more residual iconicity when compared with isolated pair developed signs, and this makes the translation from sign to meaning more transparent to naive overseers.

4. DISCUSSION

In this paper, we investigated the fitness of a small visual lexicon that evolved among members of an interacting community engaged in a graphical communication task similar to the game Pictionary. For reasons outlined earlier, the Pictionary task offers a useful vehicle to study the emergence and evolution of communication systems under controlled laboratory conditions. In particular, the corpus of signs generated by participants offers a rare opportunity to determine whether the product of cultural evolution, or glossogeny, is an optimized communication system.

The results show that graphical signs that evolve within a community offer distinct advantages when compared with those that locally develop among isolated pairs. In particular, the meaning associated with a particular sign is more accessible for a subsequent generation of sign learners (experiment 1). This decoding benefit cannot be explained by differences in speed of discriminating or ease of detecting community evolved signs (experiments 2 and 3, respectively). Instead, the benefit arises from the greater residual iconicity of community evolved signs (experiment 4). From a functional point of view, this can be explained in relation to Garrod *et al.*'s (2007) information theoretic account of different kinds of signs. They argue that icons differ from symbols in terms of where the information they convey lies. Icons work through resemblance to the objects they signify. Hence, they are effective to the extent that their graphical structure (i.e. information) maps onto the physical structure of the object. In this way, the graphical complexity of icons is related to the physical complexity of the signified object. By contrast, symbols are effective to the extent that their structure maps onto the structure of other instances of the symbol used previously to signify that object. Other things being equal, graphical symbols can become structurally simpler (i.e. bear less information) than graphical icons because they require only sufficient structure to differentiate them from other symbols in the domain. In turn, this simplification facilitates sign production making it increasingly fluent. Hence, from the point of view of communicative fitness, it makes sense for icons (or indices) to evolve into symbols. And this is what happens with both isolated pairs and in communities.

But what about the global evolution unique to the community signs? In a group context, signs need to be effective both in terms of communicative fitness within each pair of the group and in terms of transmission fitness for other group members. Our results indicate that communities achieve this by developing increasingly simple signs, but nevertheless signs that retain sufficient residual iconicity to be easily recognized (experiment 4) and learned (experiment 1) by new members of the population from which the community was drawn. So, in both community and isolated pair conditions, graphical signs evolve functionally, becoming progressively refined and therefore more efficiently produced and decoded by interlocutors. However, only community evolved signs exhibit learning and decoding benefits for persons not actively engaged in sign construction. As these benefits are unanticipated (i.e. the signs are not 'designed' with an external audience in mind), sign fitness is a 'functional by-product' of

adaptation in the community condition. Thus, like the Panamanian frog's mating call, community evolved signs are optimized in two ways at once: ease of production and ease of learning by subsequent generations. Herein lies the fitness and functionality of community evolved sign systems.

We turn now to the proximal mechanisms promoting the fitness of community evolved sign systems. Intuitively, and consistent with a biological account, the benefits of community evolved sign systems may derive from the greater pool of exemplars that communities can draw on. Community members have eight exemplars of each sign-referent pairing to select from, whereas isolated pairs have only two (i.e. one per member). While diversity is crucial to biological and cultural evolution, diversity alone cannot account for the observed benefits of community evolved sign systems. As Plotkin (2002) observed, social constructions are a product of shared agreement. Clearly, for a sign to 'work' there must be substantial, if implicit, agreement between interlocutors with regard to what the sign signifies. In other words, communication systems rely on conceptual alignment (see Pickering & Garrod 2004). For isolated pairs local alignment is sufficient, whereas for communities global alignment is necessary. We propose that the fitness of community evolved sign systems derives from the diversity of potential signs, and the need to globally align on a single sign-meaning mapping. A similar mechanism seems to operate with the development of community-wide linguistic conventions.

Garrod & Doherty (1994) examined the linguistic description schemes developed by pairs working together to navigate around a computerized maze. Three conditions were compared: isolated pairs, communities and non-communities (participants paired with a series of different partners not drawn from the same community). Like Fay *et al.* (in preparation), isolated pairs locally developed a range of different maze description schemes, whereas community members globally aligned on a single community-wide description scheme. Interestingly, the descriptions of non-community members, who were privy to a diverse range of maze descriptions, became increasingly misaligned as they encountered new partners, with participants tending to use individually salient description schemes irrespective of the scheme used by their current partner. Importantly, the description schemes adopted by community players were more efficient than those used by members of isolated pairs and non-communities, requiring fewer communicative moves to navigate successfully through the maze. The comparison between community and non-community participants indicates that diversity alone cannot account for the benefits of community evolved linguistic description schemes.

In conclusion, like the progressive adaptation characteristic of biological systems and concrete artefacts, communicative artefacts undergo a Darwinian process of survival of the fittest that promotes optimized sign systems. Via pairwise interaction, community members produce a range of competing signs, and the need to globally align on a series of sign-meaning mappings applies selection pressure. The interplay between sign diversity and this global alignment constraint results in a

series of schematized signs that retain substantial residual iconicity. This aids sign production, individual learning and the efficient translation from sign to meaning. Thus, the present study illustrates the parallels between phylogeny and glossogeny, a position consistent with a functionalist view of language. Our findings also have implications for icon design (e.g. icons for maps, computer displays, road signs and logos). The community evolved signs capture two core aspects of good icon design, concreteness and simplicity, factors that enhance individual learning and speed of processing (Gittins 1986; McDougall 2000). Harnessing the 'apparent' design prevalent in interacting communities offers an exciting alternative to traditional design practices.

REFERENCES

Barr, D. J. 2004 Establishing conventional communication systems: is common knowledge necessary? *Cognit. Sci.* **28**, 937–962. (doi:10.1016/j.cogsci.2004.07.002)

Basalla, G. 1988 *The evolution of technology.* Cambridge, UK: Cambridge University Press.

Caldwell, C. A. & Millen, A. E. 2008 Experimental models for testing hypotheses about cumulative cultural evolution. *Evol. Hum. Behav.*, 165–171. (doi:10.1016/j.evolhumbehav.2007.12.001)

Caldwell, C. A. & Millen, A. E. 2008 Studying cumulative cultural evolution in the laboratory. *Phil. Trans. R. Soc. B* **363**, 3529–3539. (doi:10.1098/rstb.2008.0133)

Christiansen, M. H. & Kirby, S. 2003 Language evolution: consensus and controversies. *Trends Cognit. Sci.* **7**, 300–307. (doi:10.1016/S1364-6613(03)00136-0)

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

Dediu, D. D. & Ladd, D. R. D. R. 2007 Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, *ASPM* and *Microcephalin. Proc. Natl Acad. Sci. USA* **104**, 10 944–10 949. (doi:10.1073/pnas.0610848104)

Fay, N., Garrod, S., Roberts, L. & Swoboda, N. In preparation. The interactive evolution of communication systems.

Galantucci, B. 2005 An experimental study of the emergence of human communication systems. *Cognit. Sci.* **29**, 737–767. (doi:10.1207/s15516709cog0000_34)

Garrod, S. & Doherty, G. 1994 Conversation, coordination and convention—an empirical investigation of how groups establish linguistic conventions. *Cognition* **53**, 181–215. (doi:10.1016/0010-0277(94)90048-5)

Garrod, S., Fay, N., Lee, J., Oberlander, J. & MacLeod, T. 2007 Foundations of representation: where might graphical symbol systems come from? *Cognit. Sci.* **31**, 961–987. (doi:10.1080/03640210701703659)

Gittins, D. 1986 Icon-based human computer interaction. *Int. J. Man-Machine Stud.* **24**, 519–543.

Goldin-Meadow, S. 2003 *The resilience of language: what gesture creation in deaf children can tell us about how children learn language.* New York, NY: Psychology Press.

Healy, G. T., Swoboda, N. & King, J. 2002 A tool for performing and analysing experiments on graphical communication. In *People and computers XVI: Proc. HCI2002: The 16th British HCI Group Annual Conference* (eds X. Faulkner, J. Finlay & F. Detienne), pp. 55–68. London, UK: Springer.

Healy, P. G. T., Swoboda, N., Umata, I. & King, J. 2007 Graphical language games: interactional constraints on representational form. *Cognit. Sci.* **31**, 285–309. (doi:10.1080/15326900701221363)

Kegl, J., Senghas, A. & Coppola, M. 1999 Creation through contact: sign language emergence and sign language change in Nicaragua. In *Language creation and language change: creolization, diachrony and development* (ed. M. DeGraff), pp. 179–237. Cambridge, MA: MIT Press.

Kirby, S. 2002 Natural language from artificial life. *Artif. Life* **8**, 185–215. (doi:10.1162/106454602320184248)

Kirby, S. & Hurford, J. 1997 Learning, culture and evolution in the origin of linguistic constraints. In *Proc. Fourth European Conference on Artificial Life* (eds S. Husbands & I. Harvey), pp. 493–502. Cambridge, MA: MIT Press.

Kirby, S. & Hurford, J. 2002 The emergence of linguistic structure: an overview of the iterated learning model. In *Simulating the evolution of language* (eds A. Cangelosi & D. Parisi), pp. 121–148. London, UK: Springer.

Kirby, S., Dowman, M. & Griffiths, T. L. 2007 Innateness and culture in the evolution of language. *Proc. Natl Acad. Sci. USA* **104**, 5241–5245. (doi:10.1073/pnas.0608222104)

Lieberman, E., Michel, J. B., Jackson, J., Tang, T. & Nowak, M. A. 2007 Quantifying the evolutionary dynamics of language. *Nature* **449**, 713–716. (doi:10.1038/nature06137)

McDougall, S. 2000 Exploring the effects of icon characteristics on user performance: the role of icon concreteness, complexity, and distinctiveness. *J. Exp. Psychol. Appl.* **6**, 291–306. (doi:10.1037/1076-898X.6.4.291)

Mesoudi, A., Whiten, A. & Laland, K. N. 2006 Towards a unified science of cultural evolution. *Behav. Brain Sci.* **29**, 329–347. (doi:10.1017/S0140525X06009083)

Pagel, M., Atkinson, Q. D. & Meade, A. 2007 Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* **449**, 717–U717. (doi:10.1038/nature06176)

Pickering, M. J. & Garrod, S. 2004 Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* **27**, 169–190. (doi:10.1017/S0140525X04000056)

Pinker, S. 1994 *The language instinct.* New York, NY: HarperCollins.

Pinker, S. & Jackendoff, R. 2005 The faculty of language: what's special about it? *Cognition* **95**, 201–236. (doi:10.1016/j.cognition.2004.08.004)

Plotkin, H. C. 2002 *The imagined world made real.* Harmondsworth, UK: Penguin.

Sandler, W., Meir, I., Padden, C. & Aronoff, M. 2005 The emergence of grammar: systematic structure in a new language. *Proc. Natl Acad. Sci. USA* **102**, 2661–2665. (doi:10.1073/pnas.0405448102)

Seyfarth, R. M. & Cheney, D. L. 2003 Signalers and receivers in animal communication. *Annu. Rev. Psychol.* **54**, 145–173. (doi:10.1146/annurev.psych.54.101601.145121)

Steels, L., Kaplan, F., McIntyre, A. & Van Looveren, J. 2002 Crucial factors in the origins of word-meaning. In *The transition to language* (ed. A. Wray), pp. 252–271. Oxford, UK: Oxford University Press.

Taylor, M. M. & Creelman, C. D. 1967 PEST: efficient estimates in probability functions. *J. Acoust. Soc. Am.* **4**, 782–787. (doi:10.1121/1.1910407)

Tomasello, M. 1999 *The cultural origins of human cognition.* Cambridge, MA: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. 2005 Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691. (doi:10.1017/S0140525X05000129)

Treutwein, B. 1995 Adaptive psychophysical procedures. *Vis. Res.* **35**, 2503–2522. (doi:10.1016/0042-6989(95)00016-X)