Available online at www.sciencedirect.com
SCIENCE DIRECT®
NEUROSCIENCE AND BIOBEHAVIORAL REVIEWS

Review

Theory of mind—evolution, ontogeny, br in mech nisms nd psychop thology

M rtin rüne*, Ute rüne-Cohrs

Received 21 April 2005; revised 2 August 2005; ccepted 8 August 2005

Abs rac

The 20c p city most highly developed in hum ns. The evolution ry origins of theory of mind c n be tr ced b ck in ext nt non-hum n prim tes; theory of mind prob bly emerged s n d ptive response to incre singly complex prim te soci l inter ction. This sophistic ted 'met cognitive' bility comes, however, t n evolution ry cost, reflected in bro d spectrum of psychop thologic l conditions. Extensive rese rch into utistic spectrum disorders h s reve led th t theory of mind m y be selectively imp ired, le ving other cognitive f culties int ct. Recent studies h ve shown th t observed deficits in theory of mind t sk perform nce re p rt of bro d r nge of symptoms in schizophreni , bipol r ffective disorder, some forms of dementi , 'psychop thy' nd in other psychi tric disorders. This rticle reviews the evolution ry psychology of theory of mind including its ontogeny nd represent tion in the centr l nervous system, nd studies of theory of mind in psychop thologic l conditions.

Theory of mind; Hum n evolution; Child development; r in mech nisms of theory of mind; Psychop thology

1. In rod c ion

The term 'theory of mind' w s origin lly proposed by prim tologists Prem ck nd Woodruff in semin l rticle to suggest th t chimp nzees m y be c p ble of inferring ment l st tes of their con-specifics (individu ls of the s me species) (Prem ck nd Woodruff, 1978). L ter on, the term w s dopted by child psychologists to describe

the ontogenetic development of mental perspective taking in infants and young children (e.g. Leslie, 1987). In terms of psychopathology, the concept of disturbed theory of mind

m tur tion (nd l rge litter size; Joffe, 1997). Prim tes, however, re extreme K-str tegists, th t is, their offspring grows slowly, multiple births re unusu l, nd birth interv ls re long. Moreover, the lre dy consider ble extension of the juvenile period in prim tes re ches m ximum in hum ns. Interesting in this reg rd is the f ct th t the length of the juvenile period in prim tes is lso positively correl ted with the size of the non-visu l cortex in the s me w y s group size is; it does not correl te with the length of gest tion, l ct tion, nd reproductive life sp n. This finding could be interpreted s supporting rel tion of slow m tur tion to constr ints of the soci l environment (Joffe, 1997). For ex mple, the extension of the juvenile period in prim tes m y h ve been cruci l to cquire the v st mount of possible soci l beh vior l 'str tegies' (procedur l rules) nd when to employ these str tegies (here, the term 'str tegy' does not necess rily imply conscious w reness; Schmitt nd Gr mmer, 1997). This process is not merely time-consuming. The re l-life opportunities of testing possible consequences of such soci l str tegies re limited in number. It is, therefore, conceiv ble th t the need for ment l simul tion of soci l inter ction might h ve speeded up the evolution of theory of mind. If ment l simul tion is involved (see below), then theory of mind not only comprises the represent tion of the ment l st tes of other individu ls, but lso one's own ment l st te (tt chment theorists h ve termed this bility 'reflective functioning'; Fon gy, 1997).

3. On ogen of heor of mind

At birth, hum n inf nts re essenti lly imm ture. The

comp red with children whose p rents use such terms less often. In ddition, the presence of older siblings speeds up young children's ppreci tion of other minds (overview in C rpend le nd Lewis, 2004). Furthermore, it is noteworthy th t, predict bly from the evolution ry fr mework outlined bove, these development l steps of theory of mind constitute hum n univers l. Although cross-cultur l evidence is still limited, Avis nd H rris (1991) h ve cle rly shown th t underst nding f lse belief emerges t simil r ge in children of the k , preliter te hunter-g therers in southe st C meroon.

Fin lly, it is noteworthy th t the development of theory of mind is cle rly p r lleled by l ngu ge cquisition. In f ct, underst nding spe ker's intention is precondition of le rning new words. As Frith nd Frith h ve pointed out, r ndom ssoci tions of utter nces with objects r rely occur when young children le rn to spe k (Frith nd Frith, 2003) nd child begins to use words undoubtedly referring to ment l st tes such s 'I think' t the ge of four—the w tershed of distinguishing between own nd other's ment l st tes.

In contr st to our growing underst nding of children's cquisition of theory of mind bilities, f irly little is known bout the development of theory of mind in dult hum ns. Due to the fund ment l role of subjective experience nd rec ll of p st soci l inter ctions in theory of mind perform nce, we would expect continuous refinement of ment l st te ttribution throughout the dult hum n life sp n. On the other h nd, selection pressure declines with ge (p rticul rly with respect to the post-reproductive life sp n). It is therefore conceiv ble th t ging does not sp re soci l cognitive bilities. Two studies of theory of mind bilities in older people h ve reve led conflicting results. H ppé et l. (1998) found th t people with me n ge of 73 ye rs, lthough slower in perform nce, were superior on theory of mind t sk comp red to dolescents nd young dults of bout 14 ye rs nd 22 ye rs of ge, respectively. In contr st, recent study reve led the opposite, n mely successive decline in theory of mind in dults ged between 60 nd 74, nd between 75 nd 89, respectively, comp red to younger dults (M ylor et l., 2002). Thus, t this st ge there is still controversy whether nd how theory of mind c p cities ch nge over the dult hum n life sp n.

4. CNS-represen a ion of heor of mind

If prim te br ins, p rticul rly neocortic l structures, enl rged over evolution ry time due to selection pressures from the soci l environment, where ex ctly is theory of mind loc ted in the hum n br in? Evidence comes from v rious sources. Comp r tive neuro n tomy nd neurophy-siology informs us which br in re s nd corresponding functions c me under selection pressure in non-hum n prim tes to evolve into the neur l correl tes of theory of mind in modern hum ns. In ddition, function l br in im ging studies nd lesion studies in p tients suffering from br in injuries or stroke m y help loc lizing the br in circuits underlying theory of mind.

efore summ rizing some of the most import nt empiric l studies, it is necess ry to point out th t divergent theoretic l conceptu liz tions of theory of mind exist. To some degree, this h s consider ble imp ct on how empiric l findings re interpreted. (1) Dr wing on Fodor's (1983) concept of modul r org niz tion of the hum n mind, some theorists dvoc te the existence of sep r te theory of mind module (e.g. Scholl nd Leslie, 1999). Like other dom in-specific cognitive c p cities represented in the br in, which process only cert in cl ss of inform tion, the theory of mind mech nism is supposed to process inform tion restricted to soci l inference. Cognitive mech nisms re ssumed to work reli bly, efficiently, nd economic lly. According to the modul r hypothesis, the development of theory of mind m inly depends on neurologic l m tur tion of the br in structures involved. Experience, on the contr ry, m y trigger the ction of the theory of mind mech nism, but does not determine the m keup of the mech nism. (2) The 'met represent tion l' theory–theory (e.g. Perner, 1991) of theory of mind is somewh t distinct from the modul r model. As Fl vell (1999) h s summ rized, the theory–theory propos l holds, simil r s the modul r theory does, th t the entities nd the c us l principles of theory of mind re s5.4(e5o)-417.9(princ7titie)-8.6(7(m y)-531..3(ntly

Now, what can we learn from primate research about theory of mind, in light of the fact that there is no unequivocal evidence of mental state attribution in non-human primates in general, and a virtual absence of theory of mind in monkeys? Single cell recordings in non-human primates convey important information about candidate cerebral representations of cognitive precursor capacities of what we call ‘true’ theory of mind in humans (the term ‘precursor capacities’ by no means ought to suggest a teleological interpretation, i.e. that something evolves in order to later suit a certain purpose).

A number of candidate structures have been identified in non-human primate brains that have undergone adaptive modifications to constitute in humans a neural network of theory of mind. Single cell recordings in macaque monkeys have revealed that neurons in the middle portion of the temporal lobe, particularly in the superior temporal sulcus (STS), selectively fire when monkeys observe the gaze direction of other monkeys. These neurons are also active when the animals observe goal-directed behavior (Gallese and Goldman, 1998). In humans, functional brain imaging studies have revealed that a homologous area of the temporal lobe is activated by observation of seemingly purposeful movements of inanimate objects (as opposed to random movements), and even when still photographs depict ‘implied’ motion (Kourtzi and Kanwisher, 2000). For example, such activity could be elicited by showing human subjects pictures of a discus thrower in action, whereas no such activity could be measured when the discus thrower was at rest. Activity in parts of the STS, therefore, is linked to the observation of intentional movements. Although this does not imply conscious awareness, the representation of ‘intentions’ is certainly a critical aspect of theory of mind. In a variety of functional imaging studies during theory of mind task performance the blood flow increased in an area of the STS adjacent to the part that was activated by monitoring biological motion (Grossman and Blake, 2002).

The temporal lobes of non-human primates also contain a specific type of cells called ‘mirror neurons’ due to their unique quality to discharge during both the execution of a certain hand or mouth action or by the mere observation of the same behavior carried out by another individual. These neurons have also been found in greater density in the ventral premotor cortex of macaque monkeys, an area that is possibly homologous to the Broca area in humans (Gallese and Goldman, 1998). In an ingenious series of experiments, the group of Rizzolatti has demonstrated that mirror neurons selectively fire when monkeys observe a hand movement of which the terminal part is hidden from their view. In other words, a subset of mirror neurons is active when the monkey can only ‘infer’ or predict the result of the incompletely visible action (Umiltà et al., 2001). Mirror neurons may therefore be crucially involved in understanding action-goal states. In humans, Fadiga et al. (1995) have shown in an experiment using transcranial magnetic stimulation (TMS) that the observation of a goal-directed hand movement elicited enhanced motor evoked potentials (MEP). Notably, these enhanced MEPs could be measured precisely in those muscles the observer would use when carrying out the action himself.

The discovery of mirror neurons in humans offers an explanation of how the ability to imitate the actions of others could have evolved into the capacity to simulate the mental states of other individuals (i.e. theory of mind) (Williams et al., 2001). However, as Frith and Frith (1999, 2001) have pointed out, for theory of mind it is not sufficient to represent goal-directed actions. It is also necessary to be able to distinguish between behavior generated by self or others. And indeed, there are at least two other important brain regions involved in theory of mind. We believe that simulating other people’s mental states does not necessarily involve conscious reflection, but is readily available to conscious awareness. For example, transference and counter-transference in dyadic psychotherapeutic settings always implicate the mutual, largely unconscious attribution of mental states such as intentions, desires and beliefs, and it is the goal of psychodynamic approaches to unveil these unconscious processes and to make them accessible to the conscious mind. For conscious reflection on one’s own and other’s mental states an individual needs computational resources beyond the capacity for imitation and action simulation, and a candidate structure involved in this task is the inferior parietal cortex. Recent research using functional brain imaging has revealed that the left and right hemisphere are differentially involved in first versus third-person perspective. First-person perspective was shown to activate the left inferior parietal cortex, whereas third-person perspective activated the corresponding region on the right side of the human brain (Ruby and Decety, 2001). Interestingly, when a subject imitates the action of another person, more activation is found in the left inferior parietal cortex, but more activation is found on the opposite side when subjects view their actions being imitated. These experimental results support the assumption that the right inferior parietal cortex may be critical for consciously representing others’ minds, whereas the left inferior parietal cortex may be involved in representing one’s own mental states (Decety and Chaminade, 2005).

The other brain area that has consistently been shown to be engaged in theory of mind is the anterior cingulated cortex (ACC). The ACC receives input from the motor cortex and the spinal cord, from the ipsilateral prefrontal cortex, and from the thalamus and brainstem nuclei (Paus, 2001). It is highly heterogeneous in terms of its cytoarchitecture and functional organization. The ACC is now conceived of as an important mediator of motor control, cognition, and arousal regulation (Paus, 2001). In monkeys, for example, the most rostral part of the ACC is active prior to the execution of self-initiated movements (Frith and Frith, 1999). Most interesting from an evolutionary viewpoint and with respect to theory of mind is that the anterior ACC inconsistently forms a paracingulate sulcus

T ble 1
Overview of br in im ging studies of theory of mind in chronologic l order

Author(s); published	S mple ( )	Me n ge	Sex m/f	r in im ging technique	ToM method/t sks	Activ ted br in re s in ToM t sks
Goel et l., 1995	9 he lthy subjects	24.7	5/5	PET [^{15}O]H_2O	Present tion of f mili r nd unf mili r objects requiring inference of others' ttribution of their function (i.e. ToM). One non-ToM condition involving inference of function of unf mili r objects from their form. Two control conditions: visu l nd sem ntic ttributes of known objects.	

Table 1 ( )

th t is present in only 30–50% of individu ls nd possibly still under selection pressure (P us, 2001). This re

to go beyond the liter l me ning of utter nces by inferring wh t the spe ker ctu lly might h ve intended (H ppé, 1994; L ngdon et l., 2002b).

In dults with psychop thologic l conditions, short stories involving double bluff, mist kes, persu sions or white lies (H ppé, 1994), c rtoons or other visu lly presented m teri l h s been used to ssess theory of mind bilities. In theory of mind rese rch in schizophreni , for inst nce, short stories with or without use of props nd picture sequencing t sks h ve been given to p tients, s well s, tests of comprehension of hints behind indirect speech, met phor nd irony. Over the ye rs, the pictori l theory of mind m teri l h s been modified in order to better control for interference with ttention, memory, 'gener l' intelligence, nd verb liz tion. One problem in e rly studies in schizophreni w s th t p tients not only performed poorly on theory of mind t sks, but lso often f iled to correctly respond to the

T ble 2

(L ngdon et l., 2001; Pickup nd Frith, 2001; overview in Frith, 2004). Th t is, these deficits re prob bly independent of other cognitive dysfunctions such s ttention, set-shifting c p city, gener l intelligence nd so forth (Lee

the results could l rgely be expl ined by this confound r ther th n by specific theory of mind deficit in AD.

y contr st, the front l v ri nt of frontotempor l dementi (fvFTD) is ch r cterized by ch nges in person l-ity nd soci l beh vior while most cognitive dom ins re rel tively preserved, t le st in the e rly st ges of the disorder. From clinic l perspective this could be indic tive of selective theory of mind deficit in FTD. In study, comp ring p tients with fvFTD with mild AD nd he lthy control subjects Gregory et l. (2002) found fvFTD p tients to perform signific ntly worse on ll theory of mind t sks with incre sing imp irment rel tive to t sk complexity. AD p tients g in f iled only on the more cognitively dem nding second order f lse belief t sks indic ting n interference with cognitive perform nce r ther th n imp ired theory of mind per se. Interestingly, theory of

rüne, M., 2005 . Emotion recognition, ‘theory of mind’ nd soci l beh vior in schizophreni . Psychi tr. Res. 133, 135–147.

rüne, M., 2005b. ‘Theory of mind’ in schizophreni : A review of the liter ture. Schizophr. ull. 31, 21–42.

rüne, M., odenstein, L., 2005. Proverb comprehension reconsidered - ’theory of mind’ nd the pr gm tic use of l ngu ge in schizophreni . Schizophr. Res. 75, 233–239.

runet, E., S rf ti, Y., H rdy- ylé, M.C., Decety, J., 2000. A PET investig tion of the ttribution of intentions with nonverb l t sk. NeuroIm ge 11, 157–166.

runet, E., S rf ti, Y., H rdy- ylé, M.C., Decety, J., 2003. Abnorm lities of br in function during nonverb l theory of mind t sk in schizophreni . Neuropsychologi 41, 1574–1582.

uitel r, J.K., v n der Wees, M., Sw b- rneveld, H., v n der G g, R.J., 1999. Theory of mind nd emotion-recognition functioning in utistic spectrum disorders nd in psychi tric control nd norm l children. Dev. Psychop th. 11, 39–58.

Grossm n, E.D., l ke, R., 2002. r in re s ctive during visu l perception of biologic l motion. Neuron 35, 1167–1175.
H ppé

Rowe, A.D., ullock, P.R., Polkey, C.E., Morris, R.G., 2001. ‘Theory of mind’ imp irments nd their rel tionship to executive functioning following front l lobe excisions. r in 124, 600–616.

Ruby, P., Decety, J., 2001. Effect of subjective perspective t king during simul tion of ction: A PET investig tion of gency. N ture Neurosci. 4, 546–550.

Russell, J., M uthner, N., Sh rpe, S., Tidswell, T., 1991. The ‘windows t sk’ s me sure of str tegic deception in pre-schoolers nd utistic subjects. r. J. Dev. Psychol. 9, 331–349.

Russell, T.A., Rubi , K., ullmore, E.T., Soni, W., Suckling, J., r mmer, M.J., Simmons, A., Willi ms, S.C., Sh rm , T., 2000. Exploring the soci l br in in schizophreni : Left prefront l under ctiv tion during ment l st te ttribution. Am. J. Psychi try 157, 2040–2042.

S ltzm n, J., Str uss, E., Hunter, M., Archib ld, S., 2000. Theory of mind nd executive functions in norm l hum n ging nd P rkinson’s dise se. J. Int. Neuropsychol. Soc. 6, 781–788.

S mson, D., Apperly, I. A., K thirg m n th n, U., Humphreys, G.W., 2005. Seeing it my w y: c se of selective deficit in inhibiting self-perspective. r in 128, 1102–1111.

S rf ti, Y., H rdy- ylé, M.C., esche, C., Widlöcher, D., 1997. Attribution of intentions to others in people with schizophreni : non-verb l explor tion with comic strip. Schizophr. Res. 25, 199–209.

S rf ti, Y., H rdy- ylé, M.C., runet, E., Widlöcher, D., 1999. Investig ting theory of mind in schizophreni : Influence of verb liz - tion in disorg nized nd non-disorg nized p tients. Schizophr. Res. 37, 183–190.

S xe, R., K nwisher, N., 2003. People thinking bout thinking people. The role of the temporo-p riet l junction in “theory of mind”. NeuroIm ge 19, 1835–1842.

S xe, R., C rey, S., K nwisher, N., 2004. Underst nding other minds: Linking development l psychology nd function l neuroim ging. Annu. Rev. Psychol. 55, 87–124.

Schmitt, A., Gr mmer, K., 1997. Soci l intelligence nd success: Don’t be too clever in order to be sm rt. In: Whiten, A., yrne, R.W. (Eds.), M chi velli n Intelligence II. Extensions nd Ev lu tions. C mbridge University Press, C mbridge, pp. 86–111.

Scholl, .J., Leslie, A., 1999. Modul rity, development nd ‘theory of mind’. Mind L ng 14, 131–153.

Sh m y-Tsoory, S.G., Tomer, R., erger, .D., Goldsher, D., Ah ron-Peretz, J., 2005. Imp ired “ ffective theory of mind” is ssoci ted with right ventromedi l prefront l d m ge. Cogn. eh v. Neurol. 18, 55–67.

Sieg l, M., C rrington, J., R del, M., 1996. Theory of mind nd pr gm tic underst nding following right hemisphere d m ge. r in L ng. 53, 40–50.

Simpson, J., Done, J., V llée-Tour nge u, F., 1998. An unre soned ppro ch: critique of rese rch on re soning nd delusions. Con. Neuropsychi t. 3, 1–20.

Simps36. M2.1ufurdR dé3TD[(ch24,)-6FostA.,