

Figure 2. Results from the cohort study. Shown are changes in L2 competence (TOEIC), GM, and FA in the cohort study; $*p < 0.05$. There were no significant differences in behavioral or imaging parameters between the TG and CG at the Pre stage. **a**, Mean total TOEIC score changes from Pre to Post-1 in CG (left of the dotted line) and from Pre to Post-1 and Post-2 in TG (right of the dotted line). Error bars indicate standard error of the mean. Asterisks show significant differences with $p < 0.05$. **b**, Right, Colored voxels represent a cluster in the right IFGop ($x, y, z = 36, 11, 28$; z value = 7.84) showing significant GM increases from Pre to Post-1 in TG compared with those in CG (group-by-time interaction, $p < 0.05$ FWE-corrected). Mean GM volume changes are shown (upper left) for the right IFGop from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. Changes in GM volume for the right IFGop significantly correlated with improvement in total TOEIC score across individuals (lower left). **c**, Right, Colored voxels represent WM clusters showing significant FA increases from Pre to Post-1 in TG than those in CG (group-by-time interaction, $p < 0.05$ FWE-corrected). Upper left, Shows mean FA changes from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. The change in FA value beneath the right IFGop was significantly correlated with change in total TOEIC score ($p < 0.05$, lower left). R, right.

Table 3. Means and SDs of subject L2 proficiency tests

	TG ($n = 24$)			CG ($n = 20$)		
	Pre	Post-1	Post-2	Pre	Post-1	Post-2
TOEIC (200)	102.5 ± 4.0	129.6 ± 5.1	114.1 ± 4.5	100.8 ± 4.3	103.7 ± 4.3	—
EVT (100)	11.3 ± 10.9	16.9 ± 8.1	12.4 ± 11.0	12.9 ± 12.1	12.0 ± 10.1	—
NART (100)	35.6 ± 4.5	37.8 ± 4.6	37.0 ± 4.0	39.8 ± 4.0	38.0 ± 4.4	—

Post-2 data, however, we noticed that three participants showed further improvement in L2 ability from Post-1 to Post-2 (8.3, 7.5, and 7.8% increases in TOEIC score), despite the overall decreases of the TOEIC score as a group. Two of those exceptional participants continued to study L2 to pass a certification examination, and one participated in a short-term study abroad program after finishing our training program. Even though these were anecdotal findings in a small number of participants, we deemed them important and performed a subgroup analysis: three participants with increased TOEIC score from Post-1 to Post-2 (continued subgroup) and the rest ($n = 21$; discontinued subgroup). There were significant differences in changes of TOEIC score from Post-1 to Post-2 between the continued and discontinued (mean $-21\% \pm 8.5$) subgroups (Mann–Whitney U test, $p < 0.001$). Intriguingly, the continued group showed increases in GM volume in the right IFGop (3.1, 2.9, and 2.5%, respectively) from Post-1 to Post-2 as opposed to the discontinued group (mean $-4.5\% \pm$ SD 1.2). The same held true for the FA value in the right sub-IFGop (2.8, 3.2, and 2.1%, respectively, for the continued group) as opposed to the discontinued group (mean $-3.6\% \pm$ 1.5; Fig. 4b). These differences reached significance for the GM change ($p = 0.001$ by Mann–Whitney U test) and for the FA change ($p = 0.001$ by Mann–Whitney U test).

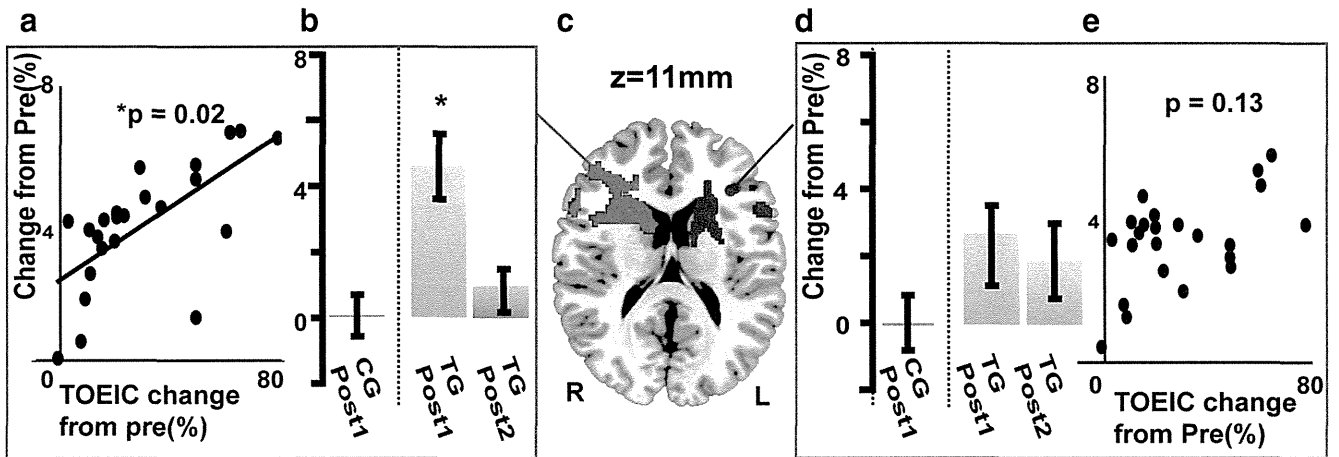
Discussion

Currently, international communication relies heavily on English. This fact imposes educational challenges to non-English native cultures in the rapidly globalizing world, prompting us to advance our understanding of the mechanisms of L2 acquisition. The human capacity for language relies on neural networks that

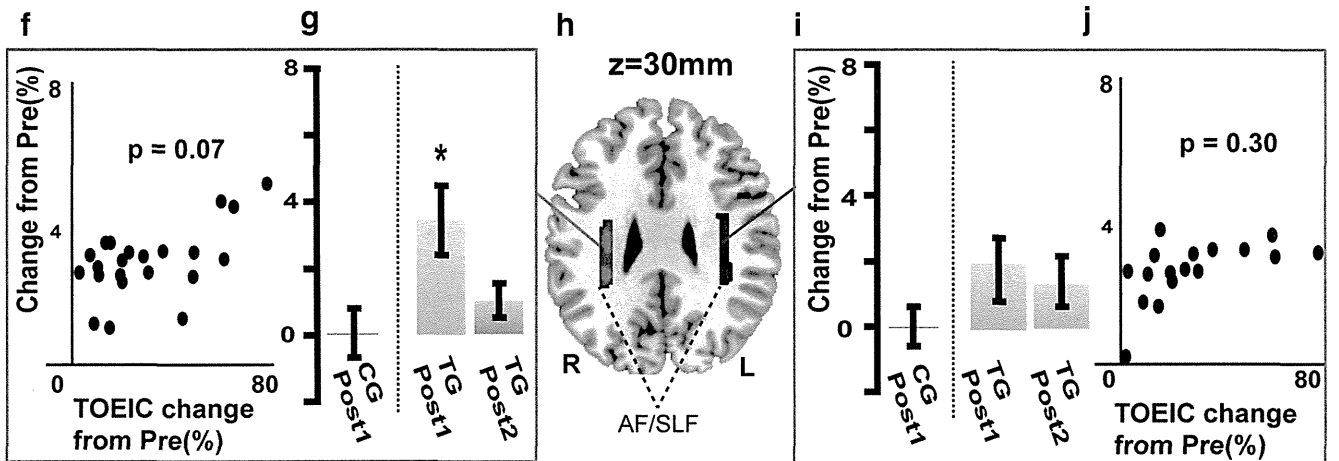
orchestrate lexicosemantic, phonological, and syntactic subsystems, which should undergo adequate tuning for L2 learning. However, nonsyntactic aspects of L2 learning are scarcely understood hitherto. Here we focused on L2 vocabulary learning, which should involve the lexicosemantic and phonological subsystems. The present cross-sectional study showed that bilateral front-subcortical-parietotemporal areas, predominantly in the right hemisphere, might underlie superior L2 vocabulary ability in late L2 learners. The cohort experiment showed that the brain architecture before learning could not predict the future gain in L2 ability after learning. This finding argues against the idea that privileged persons with developed language network have advantages in L2 learning, although it holds true for discriminating speech sounds of unfamiliar L2 (Golestani and Pallier, 2007; Golestani et al., 2007). This discrepancy suggests that relative importance of predispositions and environmental effects may differ, depending on which aspects of L2 should be learned. The cohort experiment clearly disclosed that the gains and losses of L2 ability correlated with increases and decreases, respectively, of GM structure and WM connectivity. This longitudinal consistency of the behavior–structure relationship within individuals highlights the notion that use-dependent plastic changes of neural networks underlie L2 learning. No overall gains of TOEIC score in the Post-2 suggested that the present e-training protocol was not sufficiently effective in leaving a long-term signature of L2 vocabulary learning. Note, however, that the loss of L2 ability measured by the behavioral test does not necessarily mean that the once-remembered items had completely disappeared from long-term memory. It is also possible that the items could not be retrieved in a contextually timely manner. Such presumable mechanisms agree with reduction of connectivity (i.e., disconnection) between the frontal executive areas and temporoparietal cortices.

Recent imaging studies have begun to show GM/WM changes related to learning. Biological interpretations of those imaging measures are still challenging. Many possible mechanisms could be involved: neurogenesis, gliogenesis, synaptogenesis, and vas-

IFG-CN connectivity



IFGop-STG/SMG connectivity (dorsal pathway)



IFGtri-MTG connectivity (ventral pathway)

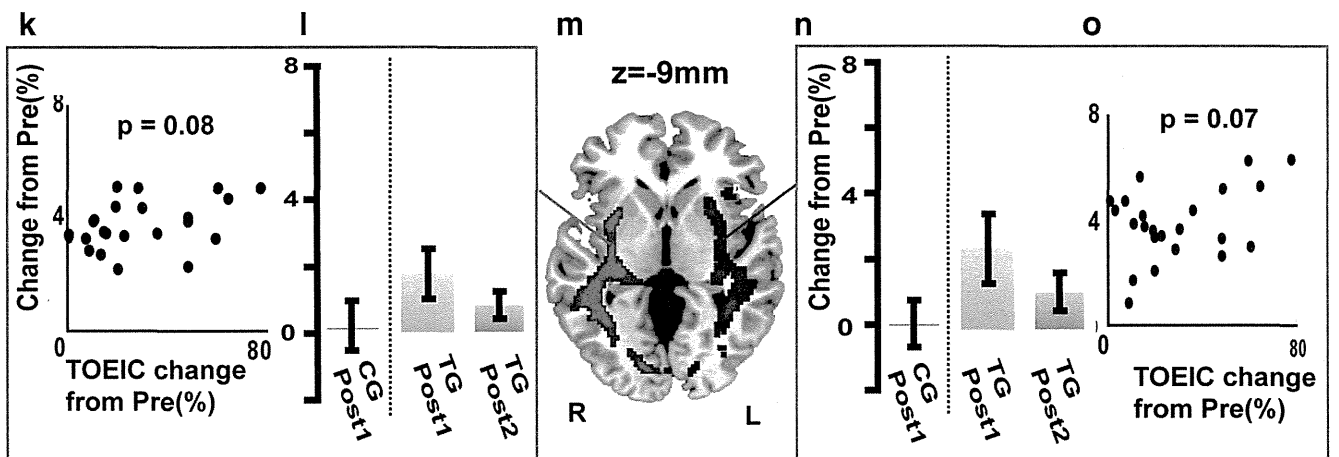


Figure 3. PDT results from cohort study. The strength of connectivity parameters for the IFGop-caudate pathway and IFGop-STG/SMG tract (dorsal pathway) were increased after the training period in the right hemisphere, but not in the left. The improvement in L2 competence (TOEIC) was correlated with the increase in the right IFGop-caudate connectivity parameter only. Error bars indicate standard error of the mean. Asterisk shows significant differences with $p < 0.05$. *a*, The change in the right IFGop-caudate connectivity parameter positively correlated with improvement in the total TOEIC score $p < 0.05$. *b*, Mean changes in right IFGop-caudate connectivity parameter from Pre to Post-1 in CG (left of the dotted line) and from Pre to Post-1 and Post-2 in TG (right of the dotted line). *c*, The red and blue areas show group-averaged IFGop-caudate tracts in the right and left hemispheres, respectively (data from Post-1 in TG). The seed and target regions were set in the right IFGop and the whole caudate nucleus (CN), respectively. *d*, Shown are left IFGop-caudate connectivity parameter changes from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *e*, The changes in the left IFGop-caudate connectivity parameter did not correlate with improvement in the total TOEIC score. *f*, Changes of connectivity parameter in the right dorsal pathway (IFGop-STG/SMG) were not significantly correlated with improvement in the total TOEIC score. *g*, Shown are changes in the right dorsal pathway connectivity parameter (*Figure legend continues.*)

cular changes for GM, and remodeling of the myelin sheath and activity-dependent axonal changes for WM reorganization indexed by FA (Scholz et al., 2009; Zatorre et al., 2012). Recently, an animal experiment associated increases in FA with increased expression of a marker of myelination (Blumenfeld-Katzir et al., 2011). An important finding here was that the bidirectional changes of the behavioral, GM, and WM parameters paralleled each other, in contrast to a previous study showing discrepancy between training-induced changes in GM and WM (Taubert et al., 2010). The coupling of behavioral, GM, and WM changes supports that learning depends upon experience-dependent activities correlated across connected regions (Fields, 2005; Zatorre et al., 2012).

The present study has provided the first compelling evidence for increased connectivity of specific tracts underlying L2 ability and acquisition. We found that the IFGop constituted networks with important language-related nodes: the STG/SMG (Price et al., 1999; Price and Crinion, 2005; Saur et al., 2008; Carreiras et al., 2009) and the caudate nucleus (Crinion et al., 2006; Friederici, 2006). The IFGop and caudate nucleus probably constitute the corticobasal ganglia circuits relevant to language processing and learning. Impairments of the caudate nucleus are identified in people with mutation of FOXP2, which is an important genetic predisposition for language acquisition (Liégeois et al., 2003; Enard et al., 2009). Additionally, caudate activity underlies lexicosemantic control in bilinguals (Crinion et al., 2006), and reduced striatal dopamine releases impair language processing (Tettamanti et al., 2005). The corticobasal ganglia circuits are suggested for reward-based reinforcement learning since they receive both contextual information from the cortex and reward signals from dopaminergic neurons (Doya, 2008). Reinforcement learning might play a part in enhancing executive control of the IFG over the mechanisms for acquiring L2 vocabulary.

We found increased connectivity of the “dorsal pathway,” corresponding mainly to the temporal part of AF involved in phonological processing (Rilling et al., 2008). AF underlies successful associative learning between sounds and words (Wong et al., 2011; Yeatman et al., 2011). AF may thus integrate phonological and se-

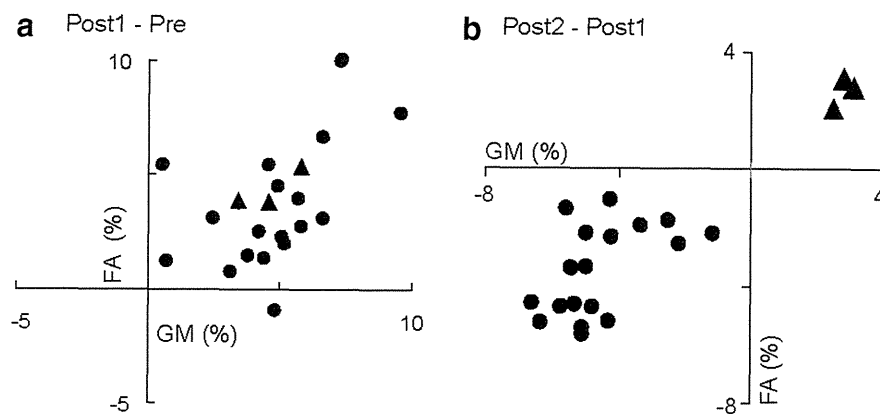


Figure 4. Correlation between GM and WM changes induced by L2 training. *a*, Changes in GM volume (abscissa) and FA (ordinate) from Pre to Post-1 in TG. The plots show significant correlation between GM volume and FA changes ($p = 0.03$). *b*, Changes in GM volume and FA from Post-1 to Post-2. The plots show significant correlation between GM volume change and FA change ($p < 0.001$, $r = 0.65$). The circles indicate participants with decreased TOEIC scores and the triangles indicate subjects with increased TOEIC scores from Post-1 to Post-2. The filled circles indicate participants who showed decrement of the total TOEIC scores from Post-1 to Post-2 while the filled triangles indicate participants who showed increment of the total TOEIC scores for the same period (see main text for the subgroup analysis).

mantic aspects of L2 vocabulary. The predominant involvement of the phonological subsystem in L2 vocabulary learning agrees with the present finding showing little connectivity changes in the ventral pathway and ILF, which are involved mainly in semantic processing (Saur et al., 2008). The predominant involvement of the dorsal pathway suggests that phonology learning may have played a pivotal role in the present L2 training paradigm. A future study, however, will be needed to dissect the substrates for the lexicosemantic and phonological aspects of L2 vocabulary learning.

A rather surprising outcome was that the IFGop-caudate-STG/SMG network correlated with L2 vocabulary competence and its learning-induced plastic changes were lateralized to the right hemisphere. These observations appear to contradict with mounting evidence indicating significance of the left hemispheric language areas for L2 (Tettamanti et al., 2002; Musso et al., 2003; Abutalebi and Green, 2007; Abutalebi et al., 2008; Sakai et al., 2009; Mårtensson et al., 2012; Schlegel et al., 2012; Ghazi Saidi et al., 2013). Moreover, learning of new languages induces plastic changes in GM and WM in the left hemisphere (Schlegel et al., 2012). In early learning stages, however, the right frontal cortex is involved in acquisition of artificial grammar and natural language (Fletcher et al., 1999, 2005; Tettamanti et al., 2002; Musso et al., 2003). Recent evidence has begun to highlight the importance of the right hemisphere for vocabulary and phonological aspects of L2 processing. The right STG/SMG is involved in non-L1 vocabulary learning (Smith et al., 2006; Jeong et al., 2010; Raboyeau et al., 2010; Veroude et al., 2010). We previously showed that right IFGop activity for switching phonology from L1 to L2 was correlated with L2 vocabulary levels (Hosoda et al., 2012). Others showed enhanced activity in the right prefrontal areas during picture naming in L2 (Videsott et al., 2010), coupling of right IFG activity with proficiency of word production in L2 (Calabrese et al., 2001; Vingerhoets et al., 2003; van Ettinger-Veenstra et al., 2010), and involvement of the right hemisphere in the control of verbal interference in bilinguals (Filippi et al., 2011). Moreover, a case study suggested an essential role of the right IFG for L2 (April and Tse, 1977); a dextral late bilingual patient suffering from a right IFG lesion had severe difficulty in finding words and reading more in L2 than in L1. Hence, evidence certainly supports the roles of the right IFG and STG in nongrammatical aspects of L2 usage, although they may not be as well recognized as the correlates of L2 compared with the left counterparts.

(Figure legend continued.) from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. Error bars indicate SEM. Asterisks show significance differences with $p < 0.05$. *h*, The red and blue areas represent the group-averaged dorsal pathway tractography including AF and SLF in the right and left hemispheres, respectively (data from Post-1 in TG). *i*, Mean changes in the left dorsal pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *j*, Changes in the left dorsal pathway connectivity parameter did not correlate with the change in the total TOEIC score. *k*, Changes in the right ventral pathway (IFGtri-MTG) parameter were not significantly correlated with improvement in the total TOEIC score. *l*, Changes in the right ventral pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *m*, The red and blue areas represent the group-averaged ventral pathway tractography in the right and left hemispheres, respectively (data from Post-1 in TG). *n*, Mean changes in the left ventral pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *o*, Changes in the left ventral pathway connectivity parameter did not correlate with improvement in the total TOEIC score. The connectivity parameters for the inferior longitudinal fascicles did not change in either hemisphere (data not shown). R, right; L, left.

A limitation of the cohort experiment was no behavioral control over CG during the training period. Although we cannot completely exclude the effects of nonlanguage e-learning factors on the training-induced changes, we considered them unlikely as the sole explanation of the learning-induced plastic changes because of the following reasons. First, there was no correlation of the total e-training time with the GM/WM changes. Second, previous long-term computer-learning studies did not report changes of right IFG, caudate nucleus, or STG (Takeuchi et al., 2010; Wan et al., 2012). Consistently, preliminary results from our computer-based sequence learning experiment for 10 weeks failed to find changes in the right IFG, caudate nucleus, or STG (C. Hosoda, M. Honda, T. Hanakawa, unpublished observation).

Paucity of evidence allows us only to speculate specific functions of the right IFG in L2 vocabulary learning. We theorize that the right IFG might link lexicosemantic-phonological knowledge between L1 and L2. For late L2 vocabulary learning, one should associate new L2 vocabulary temporarily represented in short-term memory with existing L1 vocabulary represented in the left hemisphere. This agrees with a finding that acquisition of new L2 phonology recruits the left-lateralized language network (Paulesu et al., 2009). However, more long-term encoding-retrieval processes may modify bilateral networks since the right prefrontal areas are important for accessing to long-term memory (Ranganath et al., 2007). Another simple, yet plausible, explanation is a spillover of language representations/control functions from the left to the right hemisphere. This could be particularly important here since Japanese is one of the most linguistically distant languages from English (Chiswick and Miller, 2005). The spillover concept is consistent with findings that mathematics/arithmetic experts show greater right hemispheric activity than control subjects for whom left hemispheric activity is more relevant (Hanakawa et al., 2003; Aydin et al., 2007). These findings may characterize a repertoire of use-dependent dynamic reorganization of the brain. A dramatic example is a child who shifted the originally left hemispheric language centers to the mirror-reversed, right hemispheric sites after surgical removal of the left hemisphere (Hertz-Pannier et al., 2002). Language experience-dependent changes may likely occur in the right IFG because of weaker genetic influences on the right IFG compared with the left (Thompson et al., 2001).

In conclusion, the present study indicates that the macroscopic reorganization of the IFG-caudate-STG network, especially on the right, can underlie L2 vocabulary learning in adults (1500/1500 words).

References

- Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A (2008) Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb Cortex* 18:1496–1505. Medline
- Alexander AL, Lee JE, Lazar M, Field AS (2007) Diffusion tensor imaging of the brain. *Neurotherapeutics* 4:316–329. CrossRef Medline
- April RS, Tse PC (1977) Cross aphasia in a Chinese bilingual dextral. *Arch Neurol* 34:766–770. CrossRef Medline
- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113. CrossRef Medline
- Aydin K, Ucar A, Oguz KK, Okur OO, Agayev A, Unal Z, Yilmaz S, Ozturk C (2007) Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study. *AJNR Am J Neuroradiol* 28:1859–1864. CrossRef Medline
- Beaulieu C (2002) The basis of anisotropic water diffusion in the nervous system—a technical review. *NMR Biomed* 15:435–455. CrossRef Medline
- Blumenfeld-Katzir T, Pasternak O, Dagan M, Assaf Y (2011) Diffusion MRI of structural brain plasticity induced by a learning and memory task. *PLoS One* 6:e20678. CrossRef Medline
- Briellmann RS, Mitchell LA, Waites AB, Abbott DF, Pell GS, Saling MM, Jackson GD (2003) Correlation between language organization and diffusion tensor abnormalities in refractory partial epilepsy. *Epilepsia* 44:1541–1545. CrossRef Medline
- Calabrese P, Neufeld H, Falk A, Markowitsch HJ, Müller C, Heuser L, Gehlen W, Durwen HF (2001) Wortgenerierung bei Bilingualen—eine fMRT-Studie mit Implikationen für Sprach- und Gedächtnisprozesse. *Fortschr Neurol Psychiatr* 69:42–49. CrossRef Medline
- Carreiras M, Seghier ML, Baquero S, Estévez A, Lozano A, Devlin JT, Price CJ (2009) An anatomical signature for literacy. *Nature* 461:983–986. CrossRef Medline
- Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44:953–961. CrossRef Medline
- Chiswick BR, Miller PW (2005) Linguistic distance: a quantitative measure of the distance between English and other languages. *J Multiling Multicult Dev* 26:1–11. CrossRef
- Coady J, Huckin T (1997) *Second Language Vocabulary Acquisition*. Cambridge UP.
- Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K, Usui K, Green DW, Price CJ (2006) Language control in the bilingual brain. *Science* 312:1537–1540. CrossRef Medline
- Dowens MG, Vergara M, Barber HA, Carreiras M (2010) Morphosyntactic processing in late second-language learners. *J Cogn Neurosci* 22:1870–1887. CrossRef Medline
- Doya K (2008) Modulators of decision making. *Nat Neurosci* 11:410–416. CrossRef Medline
- Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A (2004) Neuroplasticity: changes in grey matter induced by training. *Nature* 427:311–312. CrossRef Medline
- Ellis R, Heimbach R, Tanaka Y, Yamazaki A (1999) *Learning a second language through interaction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Enard W, Gehre S, Hammerschmidt K, Höltner SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, et al. (2009) A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell* 137:961–971. CrossRef Medline
- Fields RD (2005) Myelination: an overlooked mechanism of synaptic plasticity? *Neuroscientist* 11:528–531. CrossRef Medline
- Filippi R, Richardson FM, Dick F, Leech R, Green DW, Thomas MS, Price CJ (2011) The right posterior paravermis and the control of language interference. *J Neurosci* 31:10732–10740. CrossRef Medline
- Fletcher P, Büchel C, Josephs O, Friston K, Dolan R (1999) Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cereb Cortex* 9:168–178. CrossRef Medline
- Fletcher PC, Zafiris O, Frith CD, Honey RA, Corlett PR, Zilles K, Fink GR (2005) On the benefits of not trying: brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cereb Cortex* 15:1002–1015. Medline
- Flöel A, de Vries MH, Scholz J, Breitenstein C, Johansen-Berg H (2009) White matter integrity in the vicinity of Broca's area predicts grammar learning success. *Neuroimage* 47:1974–1981. CrossRef Medline
- Friederici AD (2006) What's in control of language? *Nat Neurosci* 9:991–992. CrossRef Medline
- Ghazi Saidi L, Perlberg V, Marrelec G, Péligrini-Issac M, Benali H, Ansaldo AI (2013) Functional connectivity changes in second language vocabulary learning. *Brain Lang* 124:56–65. CrossRef Medline
- Golestani N, Pallier C (2007) Anatomical correlates of foreign speech sound production. *Cereb Cortex* 17:929–934. Medline
- Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C (2007) Brain structure predicts the learning of foreign speech sounds. *Cereb Cortex* 17:575–582. Medline
- Grogan A, Green DW, Ali N, Crinion JT, Price CJ (2009) Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cereb Cortex* 19:2690–2698. CrossRef Medline
- Grogan A, Parker Jones Ö, Ali N, Crinion J, Orabona S, Mechias ML, Ramsden S, Green DW, Price CJ (2012) Structural correlates for lexical effi-

- ciency and number of languages in non-native speakers of English. *Neuropsychologia* 50:1347–1352. CrossRef Medline
- Hanakawa T, Honda M, Okada T, Fukuyama H, Shibasaki H (2003) Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *Neuroimage* 19:296–307. CrossRef Medline
- Hernandez AE, Dapretto M, Mazziotta J, Bookheimer S (2001) Language switching and language representation in Spanish-English bilinguals: an fMRI study. *Neuroimage* 14:510–520. CrossRef Medline
- Hertz-Pannier L, Chiron C, Jambaqué I, Renaux-Kieffer V, Van de Moortele PF, Delalande O, Fohlen M, Brunelle F, Le Bihan D (2002) Late plasticity for language in a child's non-dominant hemisphere: a pre- and post-surgery fMRI study. *Brain* 125:361–372. CrossRef Medline
- Hosoda C, Hanakawa T, Nariyai T, Ohno K, Honda M (2012) Neural mechanisms of language switch. *J Neurolinguistics* 25:44–61. CrossRef
- Hsu AS, Chater N, Vitányi PM (2011) The probabilistic analysis of language acquisition: theoretical, computational, and experimental analysis. *Cognition* 120:380–390. Medline
- Jeong H, Sugiura M, Sassa Y, Wakusawa K, Horie K, Sato S, Kawashima R (2010) Learning second language vocabulary: neural dissociation of situation-based learning and text-based learning. *Neuroimage* 50:802–809. CrossRef Medline
- Johansen-Berg H (2007) Structural plasticity: rewiring the brain. *Curr Biol* 17:R141–R144. CrossRef Medline
- Johansen-Berg H (2012) The future of functionally-related structural change assessment. *Neuroimage* 62:1293–1298. CrossRef Medline
- Johansen-Berg H, Baptista CS, Thomas AG (2012) Human structural plasticity at record speed. *Neuron* 73:1058–1060. CrossRef Medline
- Jou RJ, Jackowski AP, Papademetris X, Rajeevan N, Staib LH, Volkmar FR (2011) Diffusion tensor imaging in autism spectrum disorders: preliminary evidence of abnormal neural connectivity. *Aust N Z J Psychiatry* 45:153–162. CrossRef Medline
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303:1023–1026. CrossRef Medline
- Kuhl PK (2010) Brain mechanisms in early language acquisition. *Neuron* 67:713–727. CrossRef Medline
- Laufer B (1989) A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Rev* 6:10–20.
- Le Bihan D, Johansen-Berg H (2012) Diffusion MRI at 25: exploring brain tissue structure and function. *Neuroimage* 61:324–341. Medline
- Liégeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F (2003) Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat Neurosci* 6:1230–1237. CrossRef Medline
- Mårtensson J, Eriksson J, Bodammer NC, Lindgren M, Johansson M, Nyberg L, Lövdén M (2012) Growth of language-related brain areas after foreign language learning. *Neuroimage* 63:240–244. CrossRef Medline
- Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, Price CJ (2004) Neurolinguistics: structural plasticity in the bilingual brain. *Nature* 431:757. CrossRef Medline
- Meschyan GaH, A (2002) Is native-language decoding skill related to second-language learning? *J Educ Psychol* 94:14–22. CrossRef
- Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, Büchel C, Weiller C (2003) Broca's area and the language instinct. *Nat Neurosci* 6:774–781. CrossRef Medline
- Paulesu E, Vallar G, Berlinger M, Signorini M, Vitali P, Burani C, Perani D, Fazio F (2009) Supercalifragilisticexpialidocious: how the brain learns words never heard before. *Neuroimage* 45:1368–1377. CrossRef Medline
- Peltier J, Nicot B, Baroncini M, Zunon-Kipré Y, Haidara A, Havet E, Foulon P, Page C, Lejeune JP, Le Gars D (2011) [Anatomy of the periventricular white matter]. *Neurochirurgie* 57:151–155. CrossRef Medline
- Price CJ, Crinion J (2005) The latest on functional imaging studies of aphasic stroke. *Curr Opin Neurol* 18:429–434. CrossRef Medline
- Price CJ, Green DW, von Studnitz R (1999) A functional imaging study of translation and language switching. *Brain* 122:2221–2235. CrossRef Medline
- Raboyeau G, Marcotte K, Adrover-Roig D, Ansaldo AI (2010) Brain activation and lexical learning: the impact of learning phase and word type. *Neuroimage* 49:2850–2861. CrossRef Medline
- Ranganath C, Heller AS, Wilding EL (2007) Dissociable correlates of two classes of retrieval processing in prefrontal cortex. *Neuroimage* 35:1663–1673. CrossRef Medline
- Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE (2008) The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat Neurosci* 11:426–428. CrossRef Medline
- Sakai KL, Nauchi A, Tatsuno Y, Hirano K, Muraishi Y, Kimura M, Bostwick M, Yusa N (2009) Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. *Hum Brain Mapp* 30:2440–2452. CrossRef Medline
- Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008) Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A* 105:18035–18040. CrossRef Medline
- Schlegel A, Rudelson JJ, Tse P (2012) White matter structure changes as adults learn a second language. *J Cogn Neurosci* 24:1664–1670. CrossRef
- Scholz J, Klein MC, Behrens TE, Johansen-Berg H (2009) Training induces changes in white-matter architecture. *Nat Neurosci* 12:1370–1371. CrossRef Medline
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 [Suppl 1]:S208–S219. Medline
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TE (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31:1487–1505. CrossRef Medline
- Takeuchi H, Sekiguchi A, Taki Y, Yokoyama S, Yomogida Y, Komuro N, Yamanouchi T, Suzuki S, Kawashima R (2010) Training of working memory impacts structural connectivity. *J Neurosci* 30:3297–3303. CrossRef Medline
- Taubert M, Draganski B, Anwander A, Müller K, Horstmann A, Villringer A, Ragert P (2010) Dynamic properties of human brain structure: learning-related changes in cortical areas and associated fiber connections. *J Neurosci* 30:11670–11677. CrossRef Medline
- Tettamanti M, Alkadhi H, Moro A, Perani D, Kollias S, Weniger D (2002) Neural correlates for the acquisition of natural language syntax. *Neuroimage* 17:700–709. CrossRef Medline
- Tettamanti M, Moro A, Messa C, Moresco RM, Rizzo G, Carpinelli A, Matarrese M, Fazio F, Perani D (2005) Basal ganglia and language: phonology modulates dopaminergic release. *Neuroreport* 16:397–401. CrossRef Medline
- Thierry G, Wu YJ (2007) Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc Natl Acad Sci U S A* 104:12530–12535. CrossRef Medline
- Thompson PM, Cannon TD, Narr KL, van Erp T, Poutanen VP, Huttunen M, Lönqvist J, Standertskjöld-Nordenstam CG, Kaprio J, Khaledy M, Dail R, Zoumalan CI, Toga AW (2001) Genetic influences on brain structure. *Nat Neurosci* 4:1253–1258. CrossRef Medline
- van Ettinger-Veenstra HM, Ragnehed M, Hällgren M, Karlsson T, Landtblom AM, Lundberg P, Engström M (2010) Right-hemispheric brain activation correlates to language performance. *Neuroimage* 49:3481–3488. CrossRef Medline
- Van Ettinger-Veenstra H, Ragnehed M, McAllister A, Lundberg P, Engström M (2012) Right-hemispheric cortical contributions to language ability in healthy adults. *Brain Lang* 120:395–400. CrossRef Medline
- Veroude K, Norris DG, Shumskaya E, Gullberg M, Indefrey P (2010) Functional connectivity between brain regions involved in learning words of a new language. *Brain Lang* 113:21–27. CrossRef Medline
- Videsott G, Herrnberger B, Hoenic K, Schilly E, Grothe J, Wiater W, Spitzer M, Kiefer M (2010) Speaking in multiple languages: neural correlates of language proficiency in multilingual word production. *Brain Lang* 113:103–112. CrossRef Medline
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N (2011) What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *Neuroimage* 54:577–593. CrossRef Medline
- Vingerhoets G, Van Borsel J, Tesink C, van den Noort M, Deblaere K, Seurinck R, Vandemaele P, Achten E (2003) Multilingualism: an fMRI study. *Neuroimage* 20:2181–2196. CrossRef Medline
- Wan X, Takano D, Asamizuya T, Suzuki C, Ueno K, Cheng K, Ito T, Tanaka

- K (2012) Developing intuition: neural correlates of cognitive-skill learning in caudate nucleus. *J Neurosci* 32:17492–17501. Medline
- Wong PC, Chandrasekaran B, Garibaldi K, Wong PC (2011) White matter anisotropy in the ventral language pathway predicts sound-to-word learning success. *J Neurosci* 31:8780–8785. CrossRef Medline
- Yeatman JD, Dougherty RF, Rykhlevskaia E, Sherbondy AJ, Deutsch GK, Wandell BA, Ben-Shachar M (2011) Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *J Cogn Neurosci* 23:3304–3317. CrossRef Medline
- Zatorre RJ, Fields RD, Johansen-Berg H (2012) Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat Neurosci* 15:528–536. CrossRef Medline

統合失調症の社会機能をどのように測定するか

池淵 恵美

Emi Ikebuchi : How is Social Functioning of Schizophrenics Measured?

精神障害をもつ人があたり前に地域で暮らすことは重要な治療目標であるために、社会生活能力の評価は重要である。また神経・社会機能の重要性が注目され、脳機能の解明や、改善のための介入研究のアウトカムとしての社会機能の評価が重要視されるようになってきている。本論では社会機能の評価する上での視座を明確にし、現状でよく使用されている社会機能の評価尺度を簡潔に紹介し、開発が望まれる評価方法について論じ、臨床場面や研究において社会機能の評価をどう行っていくことができるかを紹介している。測定ツールを分類する基軸としては、行う能力・実世界での行動、主観的評価・客観的評価、評定・行動測定などがある。実世界での行動を評価する尺度として、NIMH-MATRICESのプロセスと並行して6つの尺度が選定されており、それぞれの特徴について概説した。ほとんどの測定方法は、情報提供者への面接に基づく客観評価である。診察室などの場でパフォーマンスを求めてその評価を行うのが「行う能力」であるが、行う能力の評価のうち、課題遂行能力については広く用いられている尺度が存在するが、対人スキルや社会的問題解決能力についてはそうした尺度はまだ存在しない。生活環境、年齢や性別やおかれている文化によって、取るべき対人行動が大きく異なることがその原因の1つであろう。診察室で測定できる脳機能と、行う能力や実世界での行動をつなぐ変数として、内発的動機づけ、メタ認知、自己効力感、得られる支援やおかれている環境などがあり、相互の関連性が低くなる要因と考えられる。臨床でアセスメントを行うには、狭義の社会機能だけではなく、こうした介入変数や、希望や支援のニーズなどの主観的評価、もともとの機能や障害の影響を知る上での生活歴が必要であり、評価尺度だけでは不十分である。介入研究の効果測定には客観的行動評定の尺度などが有用である。

<索引用語：社会機能，統合失調症，実世界での行動，行う能力，パフォーマンステスト>

はじめに

——本論の問題意識および目的——

わが国の治療文化からも医療・福祉制度の上でも、精神障害をもつ人たちがあたり前に地域で暮らすことが治療の目標になる中で、社会生活能力の評価が重要な課題になりつつある。

英国では、脱施設化が進行した1950年代に集中して評価尺度が開発された。これらの尺度は、ど

のような入院患者が地域での自立生活を送ることが可能か、また安定した生活を予測する因子は何かといった地道な研究から生まれた。WingのWard Behavior Rating Scale (WBRS)³³⁾はその先駆である。WBRSはその後、WHOの作成したPsychiatric Disability Rating Scale (DAS)³¹⁾の「病棟内行動尺度」に発展し、また地域生活を送る患者に適用できるSocial Behavior Schedule

著者所属：帝京大学医学部精神科学教室，Department of Psychiatry, Teikyo University School of Medicine

受理日：2013年3月2日

(SBS)³⁴⁾に発展した。

米国でも、脱施設化の進んだ1960～1970年代にかけて、社会機能を評価する尺度が多数開発された。それらについてはよい総説(Anthonyら¹⁾, Wallace³⁰⁾, Bakerら²⁾)がある。

わが国においても、1980年代には、多くの尺度が開発されている^{15,35,36)}。しかし、社会機能を評価する手法で、標準的に用いられる尺度はまだ見あたらないといつてよい。それは、社会機能が多岐にわたり、また社会の価値観が評価に含まれざるを得ないこと、性差、文化差、年齢などの社会的立場によって、適切な社会行動が大きく変化せざるを得ないこと、行動レベルでは把握できない評価内容も含まれることなどの原因による。社会機能という呼称に包含される概念も様々といつてよい。しかしこうした困難を超えて、標準的な社会機能の評価を確立することは、社会生活を改善する介入の標的を明確にし、効果を確定する上でゆるがせにできない喫緊の課題である。

近年はまた、神経認知機能や社会機能の重要性が注目され、脳機能の解明や、改善のための介入研究が行われるようになっており、そのアウトカムとしての社会機能の評価が重要視されるようになってきている。米国では大規模なプロジェクトとして、the National Institute of Mental Health Measurement and Treatment Research to Improve Cognition in Schizophrenia (NIMH-MATRICES)²⁰⁾が立ちあげられ、統合失調症の認知機能改善のための創薬の試みが行われている。その中で神経認知機能の評価のための神経心理テストバッテリーの標準化作業が行われた。同時に、患者や家族にとって認知機能よりも意味の大きな日常生活機能の改善に対しても評価を行い、認知機能改善薬が社会的に妥当なものであることを裏づける必要性から、MATRICSでは、もう1つのエンドポイント(co-primary measure)としての社会機能の評価についても検討が行われた¹⁰⁾。この論文でGreenらは、地域生活の実際を評価することは、環境や個人の生活経験など様々な媒介因子があつて、直接的に認知機能を反映しない可能性

があることや、認知機能の改善があつても、地域生活の改善をもたらすまでにはタイムラグがあつたり、治験では統制できない様々な要因(心理社会的リハビリテーションの有無、社会的サポート、雇用・経済情勢など)があることから、もう1つのエンドポイントとしては適切ではないと考え、認知機能の変化を直接反映するであろう評価手法として、「実際に地域でやっていることを評価する」のではなく、「どの程度行う能力をもっているか」(functional capacityまたはcompetence)と、「患者がどの程度認知機能障害を認識しているか、そしてそれがどの程度日常生活を障害していると感じているか」を評価するやり方を選択した。そして4種の評価ツールが選ばれたが、これらは神経認知機能との相関がみられる一方、実世界での行動(real-world functioning)との関連は必ずしも高くないことから、Greenらは結論として、まだ1つの尺度を選択する段階ではないとしている。行う能力(functional capacity)についても、まだ標準版はないということである。

本論においては、まず多様な社会機能を評価する上での視座を明確にし、これらの視座に沿って現状でよく使用されている社会機能の評価尺度を簡潔に紹介した上で、それぞれの評価方法の利点と限界を明らかにし、開発が望まれる評価手法について論じたい。さらに現状をふまえて、臨床場面や研究において社会機能の評価をどう行っていくかについても触れたいと考える。

1. これまでの社会機能測定ツールの現状

1. 測定ツールはどのように分類できるか

測定ツールを分類する基軸としては以下のものが挙げられる。1) どの水準の社会機能を評価するのか(行う能力・実世界での行動)、2) 誰が評価するのか(主観的評価・客観的評価)、3) どのような測定方法であるのか(評定・行動測定)、4) どのような目的をもっているのか〔治療開始前のアセスメント(機能評価)と援助プランの作成・治療の進展を把握するためのモニタリング・治療の効果を確認するための効果判定〕、5) どの

ような社会生活領域を評価するのか(居住・日常生活技能・対人技能・職業生活など).

1) どの水準の社会機能を評価するのか(行う能力・実世界での行動)

診察室などの場で与えられた社会機能の課題についてのパフォーマンスを求めてその評価を行うのが「行う能力」であり、病院の診察のための予約の電話をかける課題、物品を購入してお釣りを計算する課題、隣人と会話を続ける課題などである。「行う能力」の評価の意義についてのGreenらの見解についてはすでに前述した。それに対して、一般的には社会機能の評価でイメージされるのが「実世界での行動」の評価である。どのように情報収集を行うかで、ケアしている人へのインタビューを行う、本人に記述してもらう、日常生活の場で行動観察する、診療録などの記録をもとに評価する、これらの手段をいくつか組み合わせて行う、などの方法がある。行う能力と実世界での行動の間には、当然一定の連関があるはずであるが、先に述べた要因から必ずしも現実の評価結果においては密接な相関がみられるわけではない。この問題については、また改めて考察する。

Bowieら⁶⁾は認知機能リハビリテーションおよび社会的スキルへの介入プログラムの効果を検討するにあたり、①神経認知機能をBrief Assessment of Cognition in Schizophrenia (BACS)で評価し、②行う能力のうち対人スキルについてはthe Social Skills Performance Assessment (SSPA)を用い、課題処理スキルについてはUCSD Performance-based Skills Assessment (UPSA), The advanced finances test (小切手を書く、紙幣で支払いをする、口座に残金を残す、収支のバランスを考えるなどの課題), The Medication Management Ability Assessment (ロールプレイで複数の処方薬剤を正確な量と時間に服薬する)を用いている。そして、③実世界での行動(real-world functioning)については、the Specific Levels of Functioning Scale (SLOF)により、行動を観察して5段階判定した。このように介入を行ったときに標的となった機能とともに、

その影響がどの程度日常生活に及んでいるかを区別して評価することが、理論的には望ましい方法と思われる。

2) 誰が評価するのか(主観的評価・客観的評価)

当事者であるのか、他者であっても治療者であるのか、それ以外の専門家であるのか、家族を含めた関係者であるのかによって、得られる情報は異なる。どの評価者を選択するかは、どのような情報を得たいと思うかによって決まってくるし、どちらがどちらの代替となるという関係ではない。

3) どのような測定方法であるのか(評定・行動測定)

評定は「ほぼあてはまる、あてはまらない」などの判断を交えて評価を行うものであり、行動測定は一定の行動の生起の有無や頻度を評価するものである。前項も含めて、本人による行動報告(または自己行動監視法)や自己評定と、観察者による行動測定と評定に分けられる。容易に想像されるように、行動測定や本人による行動報告は客観性に優れ、行動療法の効果判定など、研究面で広く用いられている。Time-Sample法がその代表であろう。しかし要援助かどうかなどの判定には、社会的価値観などの介入する観察者や自己による評定が重要になってくる。具体例を挙げると、一定時間に何秒相手と視線を合わせていたかの測定は行動評価に属し、視線の合わせ方が社会的に適切かどうか、対人技能を損なう(本人が困っている)ので援助が必要かどうかの判定は、行動(自己)評定に属する。

4) どのような目的をもっているのか

入院中、デイケアなどでのリハビリテーション中、高齢者など、被験者の生活しているセッティングや状況によって、想定する機能水準が異なる。

さらに評価によって何を獲得したいかということによっても、尺度の構成は異なってくる。臨床場面で、個人の治療・援助に役立てるアセスメントにあたっては、様々な生活領域について検討を加え、介入の方策を立てることが必要で、直接の行動観察、面接、自記式調査票などが用いられる。

治療の進展度や、その後の治療目標や介入技法の修正のために行う評価がモニタリングで、自己評価法、治療場面での行動観察、家族などによる報告などが用いられる。治療効果判定にあたっては、新たな能力の獲得の程度、その能力は実際の生活でどこまで活用されているか、日常生活全般の改善度について、階層的に考える必要がある。

5) どのような社会生活領域を評価するのか

入院しているのか、仕事をしているのか、結婚して子育てしているのかなどにより、社会生活領域の重みが異なり、取り上げる評価領域が様々なことが、標準化が困難な原因の1つとなっている。実際に日常生活でどのような行動を行っているのか、そして支援の必要性がどの程度であるかを知ろうとする際には、環境評価も重要な領域である。半身麻痺の例を挙げるとわかりやすいが、どの筋に萎縮があるか、歩行が可能かどうか、近くの店まで買い物に行けるかどうかは、それぞれ関連しているもののある程度独立した事象である。ことに3点目については、車椅子が使えるか、傾斜路ができていないか、援助してくれる人がいるかなど、周囲の環境によって規定される部分が多い。

さらに社会機能の評価を複雑にする要因として、実際の社会生活では問題行動や症状などが適応に大きな影響を与えるために、それらが含まれている尺度がみられる。

2. 推奨されている社会機能の評価尺度はどのような特徴をもっているか

1) 実世界での行動を評価する尺度

Leifker ら¹⁹⁾は、NIMH-MATRICES のプロセスと並行して、実世界での社会機能の尺度策定に取り組んでいる。Validating of Everyday Real World Outcomes (VALERO) study と名づけたプロジェクトで、エキスパートによるパネルを構成し、MATRICES によって神経心理テストが選定されたのと同様のプロセスによって、既存の尺度から推奨できる「実世界での行動」評価尺度を選定する作業を行った。まず59尺度が選定され、パ

ネルによる討論を経て、社会機能（人付き合いなど）、日常生活能力（家事など）、前述の2つの混合を評価する尺度から2つずつの合計6尺度が選定された。それは以下のものである。

- ・ Quality-of-Life Scale (QLS, 混合尺度)
- ・ Specific Levels of Functioning Scale (SLOF, 混合尺度)
- ・ Social Behavior Schedule (SBS, 社会機能)
- ・ Social Functioning Scale (SFS, 社会機能)
- ・ Independent Living Skills Survey (ILSS, 日常生活能力)
- ・ Life Skills Profile (LSP, 日常生活能力)

なお VALERO では、統合失調症の当事者による自己評価は、その他の情報提供者に基づく評価との相関が低いことを指摘している。

さらに、MATRICES での議論を受けて、Kleinman ら¹⁶⁾が新たに神経認知機能を反映すると想定する社会機能評価尺度 The Schizophrenia Outcomes Functioning Interview (SOFI) を開発した。エキスパートによりまず評価の4領域が選ばれ、これまでの評価尺度を調べて362項目を抽出し、4領域に配置した。さらに122項目が追加された後で、最終的な項目を絞り込んだ。その後信頼性、妥当性、実用性などにつき、検証が行われた。そこで以上の7尺度について、1-1の項で紹介した2)~5)の項目に対応づけて表1にまとめた。なお1) どの水準の社会機能の評価するか(行う能力・実世界での行動)については、いずれの尺度も実世界での行動についての尺度であるため省いてある。

2) 行う能力を評価する尺度

前述したように、Green らは行う能力と、認知機能障害の認識を評価する尺度を検討し、以下の4種の評価ツールが選ばれた。

- ・ 行う能力の評価
 - The Maryland Assessment of Social Competence (MASC)
 - The UCSD Performance-Based Skills Assessment (UPSA)
- ・ 面接による認知機能の認識についての評価：患

表1 VALERO などで推奨されている実世界の社会機能評価尺度

*以下の項目について特徴を記載	
2)	誰が評価するのか (主観的評価・客観的評価)
3)	どのような測定方法であるのか (評定・行動測定)
4)	どのような目的をもっているのか [治療開始前のアセスメント (機能評価) と援助プランの作成・治療の進展を把握するためのモニタリング・治療の効果を明らかにするための効果判定]
5)	どのような社会生活領域を評価するのか (居住・日常生活技能・対人技能・職業生活など)
<hr/>	
●Quality-of-Life Scale (QLS) ¹²⁾	
2)	客観的評価 (患者を対象とした約 45 分間の半構造化面接)
3)	7 段階評定. たとえば「社会的発動性」については、「しばしばいっしょに何かやろうと声をかけることがありますか、それとも誰かが声をかけてくれるのを待っていることが多いですか?」「いっしょに友人と過ごすことになったときに、何をやるかは誰が決めますか?」などと質問し、6 点は全く問題なし、0 点は、活動の開始はほとんど全く他人に依存している状態
4)	地域で生活している人の欠陥症状 (deficit syndromes) をターゲットとし、主にアセスメントや効果測定で用いる
5)	全部で 21 項目. 下位尺度として、対人関係 (家事、交友、知人、社会活動、社会的ネットワーク、発動性、自閉、異性関係)、道具的役割 (職場での役割、労働能力、労働の水準、労働における満足度)、精神内界 (目的志向性、生きがい、好奇心、失快楽、目的のない活動性減少、共感、情動的交流)、公共の場での活動 (公共の場での目的、活動) がある
<hr/>	
●Specific Levels of Functioning Scale (SLOF) ²⁷⁾	
2)	客観的評価 (情報提供者による記入、または情報提供者や本人を面接して、専門家が記入)、もしくは主観的評価 (自記式)
3)	5 段階評定 (とても典型的、大体、まずまず、あまり典型的でない、ほぼ異なる)
4)	地域で生活している人のアセスメントや効果測定
5)	対人関係 7 項目、社会的に不適切な行動 6 項目、活動 (家事、買い物、金銭管理など) 11 項目、労働 (雇用されるためのスキル、最低限の指導での労働など) 6 項目
<hr/>	
●Social Behavior Schedule (SBS) ³⁴⁾	
2)	客観的評価 (本人のことをよく知っている人を半構造化面接)
3)	5 段階評定 (評価項目によってアンカリングポイントが設定されているが、問題なし、時々問題がある、ある程度問題がある、しばしば問題がある、いつも問題がある、などの段階が刻まれている)
4)	慢性の入院または施設で居住している人のアセスメントや効果測定
5)	21 領域 31 項目. コミュニケーション、対人交流、自傷行為や自殺企図、過活動や徘徊、社会的に不適切な行動、不適切な性的行動、身だしなみや衛生、行動の不活発さ、集中力など
<hr/>	
●Social Functioning Scale (SFS; MATRICS version) ⁴²⁾	
2)	客観的評価 (情報提供者の面接に基づいて評価者が記入)
3)	行動測定 (頻度や回数についての評価). 項目により何段階評価かは異なるが、領域ごとに算定基準に基づいて得点を記入する
4)	地域で社会生活を行っている人についてのアセスメント
5)	地域生活: 11 問. どういうところに住んでいるか、何時間一人で過ごすか、外出するか、自分から会話を始めるか、友人はいるか、友人や家族との程度会話するかなど. 家事: 買い物、皿洗い、入浴、洗濯、炊事、金銭管理などの頻度についての 13 問. 日常生活の質: 音楽を楽しむ、手芸、庭仕事、読書、テレビを見るなどの頻度についての 15 問. 社会的活動: 映画を見に行く、コンサートに行く、スポーツ観戦、旅行、習い事、宴会などの頻度について 22 問. 日常生活: 普段何をしているか、デイケアなどに通っていたり仕事探しなどしていないかの設問 6 問.
*職業的転帰についての尺度が別に作られている. 賃金雇用、家庭での仕事、学業の評価	

者および情報提供者への面接を実施して評価
Schizophrenia Cognition Rating Scale (SCoRS)
Clinical Global Impression of Cognition in
Schizophrenia (CGI-CogS)
これらの 4 種類の尺度は、MATRICS の 5 つの
施設で、テスト-再テスト信頼性、反復測定の可能

性、神経心理テストによる認知機能評価との相
関、実際の地域生活評価との相関、実用性・患者
の耐性が検討され、いずれも尺度としては妥当か
つ有用なものであると結論づけられたが、神経認
知機能評価との関連は UPSA が高かった。実際の
地域生活の評価との相関は、4 種ともあまり高く

表1 つづき

<ul style="list-style-type: none"> ●Independent Living Skills Survey (IJSS)³⁰⁾ * 自記式バージョンもある 2) 客観的評価 (情報提供者による尺度記入) 3) 行動測定および問題の有無についての評定, 112 項目 4) 地域での生活を送っている人 (もしくはこれから送ろうとする人) のセルフケア能力についてのアセスメント, および介入の効果判定 5) セルフケア能力の評価が目的で, 食生活, 身辺の清潔の維持, 洗濯や掃除などの家事, 身体的な健康についての管理能力, 金銭管理, 交通機関の利用法, 余暇活動, 求職や就労の能力の 8 分野にわたっている. 全体の社会適応水準も評価 * 評価の例: 「交通機関の利用 1」運転免許を所持している. 行動の頻度; (0) 全くない, またはその必要がない, (1) ときどき, (2) よくある, (3) ほとんどある, (4) いつもある. 問題の程度; (0) 問題はない, (1) 時に問題となる, (2) ときどき問題となる, (3) よく問題となる, (4) いつも問題となる
<ul style="list-style-type: none"> ●Life Skills Profile (LSP)^{11,25)} 2) 客観的評価 (専門家, 家族, 社会復帰施設のスタッフなどが質問紙に記入. 非専門家でも使用できるように工夫されている) 3) 行動評定 4) 地域での生活可能性の評定のために, 統合失調症の生活障害や, 社会生活能力の評価を目的に開発. 治療や地域での保健・福祉サービスの効果判定に有用 5) セルフケア 10 項目, 問題行動 12 項目, 社会活動 6 項目, コミュニケーション 6 項目, 自己管理能力 5 項目, 合計 39 項目について, 4 段階評定 * 評価の例: 「設問 1」被評価者は会話を始めたり応答する上で, 一般的に何か問題がありますか? 会話に全く困難はない, 若干の困難がある, ある程度の困難がある, 大きな困難がある
<ul style="list-style-type: none"> ●The Schizophrenia Outcomes Functioning Interview (SOFI)¹⁶⁾ 2) 客観的評価 (評価情報提供者と面接して, 専門家が記入) 3) 行動測定と評定が含まれる. 各機能領域について, 1~100 点までの点数を評価者がつける概括項目がある 4) 地域生活している人のアセスメントおよび効果測定 5) 居住: 4 項目. どのような形態で住んでいるのか, 居住についてどの程度の支援を受けているのかなど. 日常生活: 13 項目. 家事にどの程度の支援が必要か, 治療を順守できているかなど. 仕事: 24 項目. どのような仕事に携わっているか, 教育, 主婦など生産的活動の有無など. 社会的活動: 4 項目. 社会的な活動にどの程度参加しているか, 社会生活技能は適切かなど

なかった. その後の効果研究では, 介入の主要なアウトカム指標の1つとしてUPSAがよく利用されるようになっていく. UPSAの概観を表2に示した.

Bowieら⁵⁾は, ①神経認知機能, ②UPSAで評価した課題処理能力とロールプレイで評価する対人適応能力, ③実際の日常生活評価とを評価して, それぞれの関係についてパス解析を行っているが, 神経認知機能は, 2種類の能力(課題処理および対人適応)を介在して, もしくは直接に, 実際の日常生活に影響を与えていること, 対人適応能力は日常の対人関係と, 社会適応能力は地域生活や家事への従事と関連があるなど, 評価領域によって異なる関連性があることが報告されている.

これまで述べてきたように「行う能力」は対人機能と課題処理能力に分けて考える必要がある. 対人機能はさらに, 人と親しく付き合い信頼関係を築いていくための親和的技能と, 社会的な目的を達成するための道具的技能もしくは社会的問題解決能力に分けられる. 対人機能について「行う能力」を測定するツールは, 評価を行う場所で何らかの社会的な状況を設定し, その設定に沿って被験者と検査者とがロールプレイを行う形で行われる. 文献上よく使われている3種類の対人機能評価ツールと, 筆者が作成にかかわった改訂版ロールプレイテストとを表3に示す.

Maryland Assessment of Social Competence (MASC)³⁾は社会生活の6場面をビデオ提示し, 会話の能力を評価する. Social Skills Performance

表2 The UCSD Performance-Based Skills Assessment (UPSA)²³⁾の概要

-
- 1) 金銭管理スキル: コインや紙幣をつかって、手元にある金額を計算したり、支払いをしったりお釣りを出したりする。ガス・電気会社の支払い請求書の内容を理解しているかをみる。約5分。0~11点
 - 2) コミュニケーション: 電話によって課題を行うロールプレイ。緊急のときどこに電話するか、電話番号がわからないときはどうするかを答える。診察の予約カードをもとにして、クリニックに電話して予約を取り直すロールプレイを行う。予約カードをもとにまた受診時にまず何をやるか、受診時に何を持参するかを答える。約5分。0~12点
 - 3) 計画を立てる: 新しい水上公園についての新聞記事を読み、その理解について尋ねられた後、そこで1日過ごすために必要なものを7つ挙げることを求められる。約5分。0~14点
 - 4) 交通手段の利用: 3ルートの子バス時刻表を渡されて、特定の場所に行くにはどのバスに乗るか、料金はいくらか、交通情報を得るためにどこに電話するか、また情報が得られる場所にどうやってバスで行くか地図でも調べる。その後、約束の場所に決められた時間で行くにはどのバスに何時に乗るか、いつ着くかなどを考えてもらう。5分。0~9点
 - 5) 家事のマネジメント: デザートのレシピを渡されて、台所にある食材を点検して、何を買いに行くことが必要かを答える。5分。0~4点
-
- ・日常生活の課題についてロールプレイによって遂行することを求める
 - ・5場面
 - ・短縮版は1), 2)のみ
-

表3 「行う能力」のうちの対人機能についての評価尺度

-
- 1) Maryland Assessment of Social Competence (MASC)³⁾
 - ・社会生活の6場面をビデオ提示
 - ・設定した状況での会話の能力を評価する
 - ・言語技能 (会話の内容)、非言語的スキル (視線やジェスチャー)、全般的有効性 (話の焦点を維持しロールプレイの目標を達成する能力) の3つの軸を5段階評価

 - 2) Social Skills Performance Assessment (SSPA)²²⁾
 - ・2場面: 新しい隣人との親和的スキル、大家さんに修理を依頼する道具的スキル
 - ・ロールプレイの後で8項目を評価
 - ①シミュレーションの社会的状況に動機や興味が示しているかどうか、②会話の流暢性、③考えを明確に示しているか、④一貫して会話内容を維持できているか、⑤非言語的スキル (音調、声量、抑揚、姿勢、表情)、⑥身だしなみ (道具的スキルについては適切な問題解決が行えているか)、⑦柔軟に代替案が提示できているか、(概括的な評価)、⑧ロールプレイがその状況において社会的にみて適切でマナーを守って、相手を尊重してふるまえているか。*道具的スキルはさらに目的を達成するために一貫して交渉できているかも評価する

 - 3) Assessment of Interpersonal Problem-Solving Skills (AIPSS)⁹⁾
 - ・30秒ほどの社会的問題解決が必要なビデオ (全部で13場面) をみて、それをどうとらえるか答えた後で、相手役と問題解決を目的とするロールプレイ
 - ・どのような問題があるか、それについてどう行動するかまず設問があるが、問題が認識できていない場合にはその項目は得点なしとなる。最終的にどのような問題解決の行動をとったか、その内容、パフォーマンス、概括評価を評定する

 - 4) 改訂版ロールプレイテスト²⁶⁾
 - ・イラストによる6場面提示
 - ・ロールプレイをビデオ録画し、15項目を評定し、因子分析の結果に基づき4下位尺度としている
 - ①総合的スキル: 場面の目的の把握、対処法の起案、視線、表情、声の変化、明晰さ、流暢性、目的の達成度、社会的妥当性
 - ②スキルの主観的評価: 自己効力感と不安感の合計得点
 - ③場面の認識: 場所の認知、相手の認知 (イラストをどう把握して記憶したかについての設問)
 - ④対処法の修正

表4 本文中で用いられている略語

AIPSS	Assessment of Interpersonal Problem-Solving Skills
CASIG	Client Assessment of Strengths, Interests, and Goals
CGI-CogS	Clinical Global Impression of Cognition in Schizophrenia
DAS	Disability Rating Scale
ICF	International Classification of Functioning, disability and health
ILSS	Independent Living Skills Survey
LSP	Life Skills Profile
MASC	Maryland Assessment of Social Competence
NIMH-MATRICES	National Institute of Mental Health Measurement and Treatment Research to Improve Cognition in Schizophrenia
QLS	Quality-of-Life Scale
SBS	Social Behavior Schedule
SCoRS	Schizophrenia Cognition Rating Scale
SFS	Social Functioning Scale
SLOF	Specific Levels of Functioning Scale
SOCI	Staff Observations and Client Information
SOFI	Schizophrenia Outcomes Functioning Interview
SSPA	Social Skills Performance Assessment
UPSA	UCSD Performance-Based Skills Assessment
VALERO	Validating of Everyday Real World Outcomes (VALERO) study
VRFSa	Virtual Reality Functional Skills Assessment
WBRs	Ward Behavior Rating Scale

Assessment (SSPA)²²⁾は2場面において親和的技能和道具的スキルを評価する。Assessment of Interpersonal Problem-Solving Skills (AIPSS)⁹⁾は社会的問題解決が必要な13場面について評価する。改訂版ロールプレイテスト²⁶⁾は受信・処理・送信技能の3段階を評価できるのが特色である。

なお本論においては評価尺度の様々な略語が出現するので、読者の便宜のためその一覧を表4に掲げる。

3. 評価水準の違いに基づく特徴

1) 実世界での行動の評価

表1で示した7尺度ともに客観的評価であり、

実世界での社会機能を測定するためにはやはり客観的評価がよいことがわかる。主観的評価は、本人が生活の質をどう評価するかを知り、援助を組み立てていく上で抜きがたい重要な情報である。

測定の方法の上では、情報提供者への面接が多く、情報提供者による尺度記入もみられる。評価者自身がつぶさに実際の生活を観察することは実際には困難であるので、care giverによる情報の確度が高く実用的であるためと思われる。より確度を高めるために、面接と並行してcare giverの記録を参照したり、複数の情報提供者にあたる方法をとっている尺度もあるが、簡便性に欠け実施に困難があるために多くは採用されていない。面接の結果、特定の行動について評定を行って尺度を記入するやり方が多く、もしくは行動測定と評定とを併用しているものもある。たとえば「洗濯を週2回以上している」という情報を記載するのは行動測定であるが、それが果たして被評価者にとって適切であるのかどうかの判断が入らないと、臨床的に有用な情報とはならない。そのため表1に取り上げた尺度はいずれも行動評定を取り入れている。評定を行う判断が信頼性を確保するうえでは、丁寧なアンカリングポイントが設定されていなければならないだろう。また情報提供者によって1つ1つの行動を取り上げるのではなく、家事なら家事といった特定の機能領域について概括的な評定を行うやり方は簡便である。たとえばLSPはそれぞれの領域の行動群に対しての概括的評定が求められるために項目数が少なく、評価の所要時間も10分程度である。しかし評定は、地域や文化や支援の質によって、「妥当であるか」「援助が必要であるか」が異なってくるという限界を含んでいる。一方ILSSでは個別の行動それぞれについて行動測定を求めており、項目数が多く評価に時間がかかる反面、より客観性の高いものとなっている。ツールに何を求めるかによって、とられる方法論は違ってくることになる。

取り上げられる評価領域はどのようなpopulationを対象とするかによって適する尺度が異なる。たとえば入院患者であればSBSが適用となる

だろうし、地域で一人暮らしがどの程度やれているかをみるためにはLSPが役立つだろうし、仕事についている人などより社会機能の高い人についての評価はSFSが有用である。社会機能は初めにも触れたように広汎であり、機能水準にもずいぶん幅があるために、1つの尺度でカバーしがたいように思われる。筆者は以前、どのような社会機能測定ツールが最もよく使われているか、発表されている論文から調査してみたことがあるが、Global Assessment Scale (GAS) もしくは Global Assessment of Functioning (GAF) が最もよく使われていた。それは簡便性に加え、1つの尺度で全部の機能水準をカバーできる利便性があるからだと思われる。ただし評価者の主観的評価であるがために信頼性に欠ける危険性についても、留意が必要である。

2) 行う能力の評価

行う能力についてはパフォーマンステストを行うことになるが、限られた時間や診察室などでの実行可能性の枠組みの中で、「どのような社会的場面を設定するか」が大きな主題となる。たとえば、自立生活をする上でどの程度一人で「行う能力」をもっているかを評価するなど、目的となる社会生活を限定すれば、ある程度選ぶべき社会的場面を設定することができる。しかし、行う能力が不十分な場合に、それまでの体験などが影響している可能性は考慮する必要があるだろう。普段調理などしたことがないものに対して、献立が与えられて必要な食材を選ぶ作業は、能力そのものというよりは、被験者の過去の体験や社会的知識が大きく影響するからである。

対人機能についてはより複雑で、課題処理能力で述べたどのような生活環境かだけではなく、被験者の年齢や性別やおかれている文化によって、とるべき対人行動が大きく異なってくるので、「よく遭遇する社会的場面」を設定することはかなり困難である。そのために12場面など、多くの社会生活場面を取り上げているテストもあるが(表3)、実施に時間がかかる。

もう1つの主題は、対人機能の評価方法の問題

である。表3に示したように、既存のロールプレイテストはいずれも、視線や表情などの外見の行動測定を踏まえつつ、とられた対人行動が適切なものであるのかを評定する構造となっている。しかし対人行動を記述する変数は多数あり、その重みづけが明らかでなく、さらに対人行動をすべて数量的に記載できるとは言い難い。アウトプット(表情、明晰さ、声の変化など)は相互の関連性が高く(弁別性が悪く?)¹³⁾、個々に評価する意義については疑問がある。ヒト以外の動物に対してはある程度行動測定で記述できるかもしれないが、ヒトの場合には言語がコミュニケーションに占める割合が大きい。既存のツールでは、評価者による価値判断を挟むことにより、複雑な対人行動を縮約して概括的な評価を行うとしており、評価を行うにあたって、社会的状況を念頭において、評価者の主観に頼る部分がかかなり残っているからこそ、状況把握、妥当性や行動の効果について、評価できている面がある。このやり方は実用的であるものの、深刻な信頼性の問題をはらんでいる。性別や年齢や文化差についても、筆者がかつてカリフォルニアでのロールプレイに参加したときに、学校で注意を受けた生徒はまっすぐ教師の目を見るよう指導されていた。日本であったなら、うなだれている生徒の方がより反省していると評価されるかもしれない。あやまるときの適切な態度は当然のことながら女性と男性では違うし、年少のものと年配のものとは違うだろう。こうして人間の社会的な価値観を暗黙の前提とした評定によって、行われたパフォーマンスを評価しているというのが現状と思われる。

3つ目の主題は、社会的場面でのダイナミックな相互作用をどの程度評価に含むことができるかということである。会話を例にとるとわかりやすいが、話し手と聞き手は、受信-処理-送信機能を繰り返して1つ1つの会話を紡いでいく。相手はどう反応するかで次の言葉が違ってきて、同じテーマで始めた会話も全く異なる展開になる可能性がある。しかしパフォーマンステストでは、テストの再現性の制約から、1つの社会的場面を提

示して、それにどう反応するかという、一往復限りの短い刺激提示にせざるを得ないし、何回か会話をするにしても、信頼性の制約から、相手役の反応もあらかじめ決められたものにせざるを得ない。ダイナミックな相互作用をみることはできない(みないことにしている)設定なのである。こうしたことから、対人機能の評価も、表3で示したようにアウトプットとしての送信技能にほぼ偏っており、受信-処理技能も含む評価になっているのは改訂版ロールプレイテストのみとあってよいと思われる。

4つ目の主題は、認知機能や実世界での行動との関連性である。Ikebuchiら¹⁴⁾が示したように、行う能力のうちの対人機能は精神症状や特定の神経認知機能や社会的認知との関連性がみられる一方で、前述のように実世界で行動との関連性は高くない。それについてはこれまでの3つの主題で述べてきた、既存のパフォーマンステストの限界であるとともに、次項で述べる種々の要因が関与していると思われる。

II. なぜ測定ツールと実世界での機能との懸隔があるのか

1. 脳機能の直接的な測定と実世界での機能との間に横たわるもの

診察室で測定できる脳機能と、行う能力や実世界での行動との懸隔をつなぐ変数として、近年いくつかが注目されてきているものがある。

まずは内発的動機づけである。実験社会心理学の概念で、金銭や食べ物、名誉などのような外的報酬に基づく外発的動機づけと対比され、心の中の満足感を得ることを目的としており、近年脳機能の解明が行われている。精神症状としての意欲の低下とも関連があり、主観的な体験からいえば、希望や将来の展望とかかわりがある。内発的動機づけについてはChoiらの開発した尺度(Intrinsic Motivation Inventory)⁷⁾があり、興味、価値観、努力、緊張、選択性の5下位尺度・21項目の自記式尺度である。この5下位尺度からも内発的動機づけの構成要因を推測することができる。

次にメタ認知がある。現在進行中の自分の思考や行動そのものを対象化して認識することにより、自分自身の認知行動を把握することができる能力であり、「知っているということを知っている」「認知していることを認知している」などである。認識の対象としては、認知機能、自己の心理的状态、精神障害によって起こった変化や日常生活の障害などがある。環境に合わせてスキルを調整する能力とかかわりがあり、セルフモニタリング能力とも関連がある。精神療法の文脈では、メタ認知として自分の心の理解、他者の心の理解、他者が独立した存在であることの理解、自身の考えや精神状態についての洞察(mastery)などが想定されている。

3番目の介在変数として、実際に行動を起こす上でどの程度うまくやれると考えているかという自己効力感(self-efficacy)や、自分ほどの程度行う能力があるかという判断である「自覚している能力(perceived competency)」や、行動を起こす上で感じる不安がある。これらの項目には相互に関連があることが推測できる。また内発的動機づけと自覚している能力には相関があることが、Choiら⁸⁾、Tasら²⁹⁾の報告でわかっている。自己の能力についての認識や感じる不安と、メタ認知との間に関連性があることも容易に推測できる。以上のように、内発的動機づけ、メタ認知、自覚している能力などの3種類の介在変数はそれぞれ独立ではないと考えられる。

4番目には、実際の能力が発揮できるかどうかについては、おかれている環境や支援が得られるかどうかという面も大きい。これは半身麻痺の例を挙げて前述したが、行う能力を発揮するパフォーマンスに教示や手がかりや安心できる保護的な環境が与える影響はあるだろうが、おそらく大きな影響を与えるのは行う能力を実世界で実行する場合においてであろう。

図1にこれまで述べてきた介在変数と測定される社会機能との関係を簡単なシェーマにして示した。3番目の介在変数についてはとりあえず自覚している能力を取り上げてある。図1では実世界

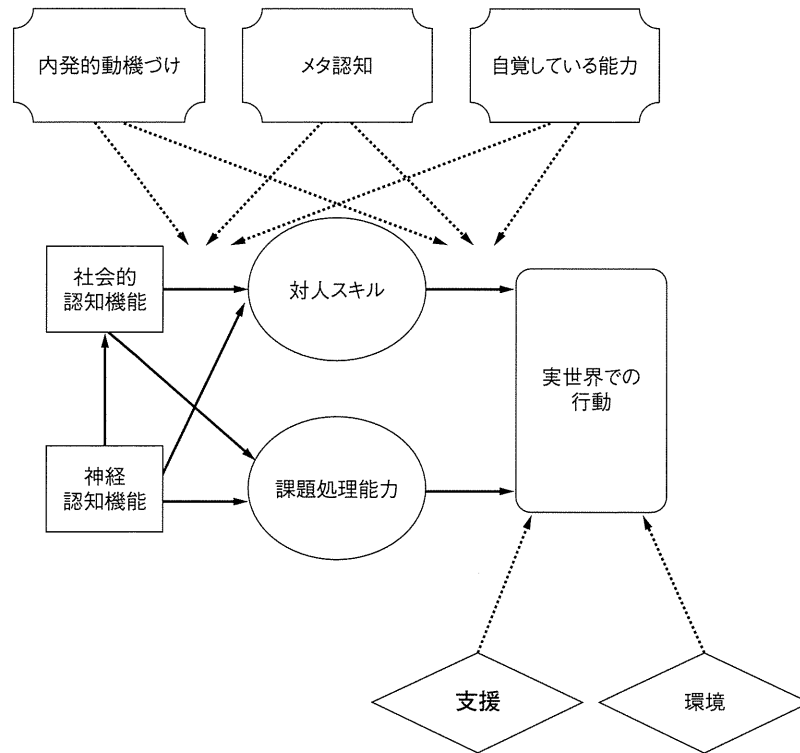


図1 社会機能に影響を与える要因

- ・ 行う能力 (functional capacity, competence)：対人スキル，課題処理能力
- ・ 実世界での行動：セルフケア，日常生活技能，社会での対人交流，就労・就学生活など

での行動として表現しているが、測定ツールで評価している評価内容と実世界での実際の行動には当然距離があるだろう。測定方法によりどのような違いが生じるか異なるので、評価結果を解釈するにはどのような誤差が生じやすいかについての知識が望まれる。同様にして実体としての行う能力と、測定ツールで測定された「行う能力」の間にも距離があるだろう。

私たちが測定ツールを使って社会機能を測定する際には、図1のどの部分を実際は測定しているのか常に明確にしておく必要があるし、測定と実態との誤差についても念頭におく必要があるし、想定されたものを実世界へと拡張する際には、介入変数を計算に入れることが必要になってくるだろう。さらには実際の生活の上では、支援を得たり代償するスキルをもっていることで、大きな困

難は回避できる場合もあり、できること・やっていることだけでは支援のニーズは測れないこともありうるだろう。

2. どのような社会機能の測定技術を発展させるべきか

1) 実世界での行動の測定

繰り返し述べているように社会生活の多様さが測定ツールの標準化を阻んでいる。特定の文化のもとで作られた測定ツールは他の文化のもとではうまく機能しないのが現状である。普遍的な社会機能が概念として抽出され、その文化差によるあらわれ方が把握できるのであれば、普遍的な評価尺度とその各文化への適合版が策定できるだろうが、現状ではそれは困難である。

実際の行動を価値判断抜きで測定して記載し、

さらにそれを各生活状況のもとで適応的であるか・支援の必要があるかという価値判断を加えていくという方法も、標準化のやり方としては考えられる。実際の行動を記述する目的で作成されたのが、改訂版国際障害分類 (International Classification of Functioning, disability and health: ICF)³²⁾で、身体構造と身体器官系の生理的機能 (脳機能を含む機能障害)、活動 (個人による課題または行動の遂行)、参加 (社会的な生活状況) の3層構造を想定して評価する。そして環境による影響を加味して、標準的環境もしくは最も促進的な環境で示す最高能力と、現実の環境における実際の生活能力とを分けて記載することになっている。活動および参加の評価は9領域あり、学習と知識の応用、一般的な課題と遂行要求、コミュニケーション、運動能力、セルフケア、家庭生活、対人関係、主要場面での生活、地域・社会・市民生活となっている。ICFは実世界での行動について、正確に記述・分類する上で有用である。一方では、評価に時間がかかること、望ましい行動かどうかといった価値判断を含まないがために、臨床場面では使い勝手が悪い印象が残る。

以上のように、まだ標準版策定の方法論が打ち出せないように筆者は感じている。VALERO studyで試みられたように、各文化で適切な測定ツールを推奨し、それぞれの互換可能性について検討しておくことが1つの解決法となる可能性がある。そうであるならば、わが国においても推奨する測定ツールを選定する作業が必要になるだろう。環境や支援の影響による修飾については、ICFがとったように、標準的な環境と現実の環境とで分けて記述するやり方があるだろう。

2) 行う能力の測定

実用性の観点からは、実際の実施時間は1時間以内で可能であることが望まれるので、パフォーマンスを求める際の刺激場面の数が限定されるし、社会生活を代表する場面を用意する必要がある。ここでも標準的な社会的な生活場面とは何かという問題が生じてくる。ヒトとして社会での集団を形成し、親和的な関係を維持したり社会生活上

の課題を達成したりする上で、本質的な基本スキルとはなんだろうか。霊長類などとの比較や発達の見点からそうした基本スキルが明らかになってくることを期待したい。

刺激提示の方法についても、どのようなやり方が望ましいか課題が残る。文章やイラストによる提示は社会的認知の一部しか刺激しないし、ビデオによる提示はその画面に入り込む想像力が要求されるし、社会的文脈を教示文などの形で補う必要があるだろう。Virtual Reality Functional Skills Assessment (VRFSA)²⁴⁾は Head mount display を装着し、6場面 (初めての人に自己紹介、友人と約束するなど) でアバターが現れて話しかける。virtual reality は実世界に近い体験が得られるものの、このVRFSAはアバターとのやりとりはできず、相互交渉をするところまでは至らない。反応連鎖的な virtual reality の手法はすでにパソコンゲームなどでは取り入れられているものの、自在な会話というところまではいっていない。生身の人間が現れて実際に社会的場面を創出する方法は最も実世界に近いだろうが、再現可能性や実施する手間の点で、実用性にやや欠ける。結局のところ、virtual reality の技術の発展が今後の1つの方向性だろうか。

評価の変数をどう選択し、しかも実世界の重要性に沿って、多数の測定変数の統合と重みづけをしていくにはどうしたらよいただろうか。たとえば先ほどのVRFSAでは、会話の開始までの時間、アバターとの距離間隔、視線の合わせ方、会話の長さを選定しているが、いずれも自動的・客観的に測定しやすい変数ではあるが、これらの変数のみで会話の質をはかることは妥当とは思えない。言語の表出でも、音声の質やピッチ、大きさなどの変数があるが、会話内容をどう数量化できるだろうか。

複雑な社会的行動を記述するには、多数の変数について評価せざるを得ないだろうが、解析方法についても実用性や妥当性の点で、さらに開発の余地がある。Koshibaら¹⁷⁾は、Bouquet法 (Behavior Output Analysis for Quantification of Emo-

tional State Translation) により、家禽ヒヨコ、マーモセットでの経験をもとに、アスペルガー障害児の社会的行動の質を判別することを行っている。具体的には社会的行動をビデオ撮影し、移動速度、視野の向き、他個体との距離などを求める。算出したパラメータ m 個の m 次元ベクトルの時系列情報のデータを主成分分析 (PCA) し、特徴空間上のベクトルに変換し、第1成分と第2成分の2次元にその結果をプロットすることで、時間変化を加えた3次元で経時曲線が得られる。行動測定だけではなく、血中ホルモン濃度などの検査所見、診断スコアや症状評価などの離散値をもつスコア、赤外線画像、脳波などのデータも標準化して解析に加えることができる。こうした手法により、即時に多数の変数を統合したデータを表示することが可能なので、やはり妥当な変数の選定とアンカリングポイントを明確にした評定の信頼性の向上が当面の課題となるだろう。

Ⅲ. 社会機能を標的とした介入研究および 日常の臨床でどのように測定ツールを 用いることができるか

すでに述べたように、社会機能を測定する目的には、アセスメントとモニタリングと効果測定がある。アセスメントでは狭義の社会機能だけではなく、図1に示した介入変数についても考慮する必要がある。臨床場面で「本人がどのような生活を希望しているのか」「どのような支援のニーズをもっているのか」が重要であることと、これは符合している。アセスメントの手続きを構造化したものとしては、ケアアセスメント票²⁸⁾や Client Assessment of Strengths, Interests, and Goals (CASIG)¹⁸⁾がある。ケアアセスメント票は主に地域で生活する、保健福祉などのケアサービスを必要とする人に対してケアマネジメントを実施する際のアセスメントとケア計画立案に用いる。本人の希望、ケアの必要度、環境条件、本人がかかえる社会生活上の困難な問題、ニーズのまとめとケア目標の5部門により、本人や関係者との面接、家庭訪問などによる行動観察、医療録やこれまで

の援助記録、関係者からの情報に基づいて評価する。CASIGはケアアセスメント票と同様の対象者が想定され、入院患者用と、地域生活者用の2タイプがある。医療・保健・福祉サービスの専門家や従事者が、サービス計画を立案したり、本人の目標を明らかにしたり、提供されるサービスをモニターしたりすることに用いる。本人との60~90分の面接が中心で、家族や関係者からも情報収集する。本人の希望、志向、強みを重視して評価するのが原則である。専門家による評価尺度 (Staff Observations and Client Information : SOCI) と組み合わせて用いる。SOCIはCASIGと評価項目が全く同一で、両者をつきあわせた上で、本人と専門家とで面接を行いリハビリテーション計画を立てるよう作成されている。

ケアアセスメント票やCASIGは包括的で支援の方向が導かれるように工夫されているものの、支援の策定は実際にどのような支援が提供できるかによって大きく影響されるし、支援を行いつつ次の支援を導いていくというらせん式の展開が臨床場面ではしばしば一般的に用いられるために、必ずしも普及していないのが現状であろう。実際の臨床場面で行われているエキスパートのアセスメント方法について、構造化する工夫が求められているとも言えるかもしれない。さらに支援の成否には、現在の能力だけではなく、過去の生活全般にわたってどのような社会生活能力を發揮してきたのか、もともと発達が不十分であったのか、疾患によって能力低下が生じているのかを知ることが有用であり、通常はこれまでの診療・援助記録や、当事者や家族との面接によって情報を得ることになる。数量的な社会生活評価尺度だけでは代替できないといえるだろう。

モニタリングは支援の結果標的とした機能がどのように変化したかを知るために行うので、特定の機能に的を絞った行動測定が役立つ。また自己評定法や自己行動観察法が有用であろう。

効果測定において、本論で紹介してきた社会生活評価尺度が最も利用されるだろう。実世界での行動については、世界標準版尺度の選定作業と連

動しつつ、それをわが国の文化や環境に適合した尺度を選定していく作業が求められるだろう。現状でいえば、VALEROなどで推奨されている尺度群がその候補であり、評価の方法論により、また取り上げる評価領域によって、使い勝手が異なるので、使用目的によって複数の尺度を使い分けるのが現実的かもしれない。行う能力、ことに対人技能の評価については、まだまだ課題が大きく、標準版策定までは距離がある。

おわりに

以上をまとめると、社会機能のアセスメントにおいては、ケアアセスメント票やCASIGのような、行動測定・評定とともに当事者や家族の希望を含む主観的評価、提供できる支援などの環境評価など、総合的なアセスメントシステムが開発されている。さらに臨床現場での使い勝手や有効性を上げていくためには、精神科医を含め経験を積んだ社会生活支援の専門家が行っている多面的なアセスメントを整理して明文化し、初心者にも理解できる形で整理していくことが有望ではないかと思われる。社会機能のアセスメントに限らず、初診時の見立てと治療の開始など、専門家・当事者・環境との複雑な相互作用の中で、適切に情報収集し、その上で意思決定していくことは、熟練した専門家がartとして行っていることであるが、それをアルゴリズムなどの形で明文化していく技術が求められている。モニタリングについてはすでに述べたように、介入の標的を明確にした上で絞った認知・行動について測定や評定を行っていくことになるので、そのためのわかりやすいモニタリングシートなどの開発が有用であろう。社会機能への治療や支援効果の測定については、実世界での行動については本論の表1に示した尺度の中から、被験者の生活状況や介入の目的に合わせて選定していくことが現段階では実際的と思われる。その上で標準化作業の道筋を議論していくことになるだろう。行う能力については、当面は表2や表3の尺度を利用していくが、本論で様々な課題があることを論じているので、それ

を土台として新たな方法論が開発されていくことが望まれる。今後のこの分野の発展に期待したい。

なお、本発表に関連して開示すべき利益相反はない。

文 献

- 1) Anthony, W., Cohen, M., Farkas, M.: Psychiatric Rehabilitation (高橋 亨, 浅井邦彦, 高橋真美子訳: 精神科リハビリテーション. マイン, 神奈川, 1993)
- 2) Baker, R., Hall, J.N.: REHAB A new assessment instrument for chronic psychiatric patients. *Schizophr Bull*, 14; 97-111, 1988
- 3) Bellack, A.S., Brown, C.H., Thomas-Lohrman, S.: Psychometric characteristics of role-play assessments of social skill in schizophrenia. *Behav Ther*, 37; 339-352, 2006
- 4) Birchwood, M., Smith, J., Cochrane, R., et al.: The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programs with schizophrenic patients. *Br J Psychiatry*, 157; 853-859, 1990
- 5) Bowie, C.R., Depp, C., McGrath, J.A., et al.: Prediction of real-world functional disability in chronic mental disorders: A comparison of schizophrenia and bipolar disorder. *Am J Psychiatry*, 167; 1116-1124, 2010
- 6) Bowie, C.R., McGurk, S.R., Mautsach, B., et al.: Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. *Am J Psychiatry*, 169; 710-718, 2012
- 7) Choi, K.H., Mogami, T., Medalia, A.: Intrinsic motivation inventory: an adapted measure for schizophrenia research. *Schizophr Bull*, 36; 966-976, 2010
- 8) Choi, K.H., Saperstein, A.M., Medalia, A.: The relationship of trait to state motivation: The role of self-competency beliefs. *Schizophr Res*, 139; 73-77, 2012
- 9) Donahoe, C.P., Carter, M.J., Bloem, W.D., et al.: Assessment of interpersonal problem-solving skills. *Psychiatry*, 53; 329-339, 1990
- 10) Green, M.F., Nuechterlein, K.H., Kern, R.S., et al.: Functional co-primary measures for clinical trials in schizophrenia; Results from the MATRICS psychometric and standardization study. *Am J Psychiatry*, 165; 221-