Self-supervised, mobile-application based cognitive training of auditory attention: A behavioral and fMRI evaluation

Josef J. Bless [a,b,*], René Westerhausen [a,b], Kristiina Kompus [a,b], Magne Gudmundsen [a], Kenneth Hugdahl [a,b,c,d]

[a] Department of Biological and Medical Psychology, University of Bergen, Norway
[b] NORMENT Center of Excellence, University of Oslo, Norway
[c] Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
[d] Department of Radiology, Haukeland University Hospital, Bergen, Norway

ARTICLE INFO

ABSTRACT

Emerging evidence of the validity of collecting data in natural settings using smartphone applications has opened new possibilities for psychological assessment, treatment, and research. In this study we explored the feasibility and effectiveness of using a mobile application for self-supervised training of auditory attention. In addition, we investigated the neural underpinnings of the training procedure with functional magnetic resonance imaging (fMRI), as well as possible transfer effects to untrained cognitive interference tasks. Subjects in the training group performed the training task on an iPod touch two times a day (morning/evening) for three weeks; subjects in the control group received no training, but were tested at the same time interval as the training group. Behavioral responses were measured before and after the training period in both groups, together with measures of task-related neural activations by fMRI. The results showed an expected performance increase after training that corresponded to activation decreases in brain regions associated with selective auditory processing (left posterior temporal gyrus) and executive functions (right middle frontal gyrus), indicating more efficient processing in task-related neural networks after training. Our study suggests that cognitive training delivered via mobile applications is feasible and improves the ability to focus attention with corresponding effects on neural plasticity. Future research should focus on the clinical benefits of mobile cognitive training. Limitations of the study are discussed including reduced experimental control and lack of transfer effects.

1. Introduction

In health and sickness, humans are preoccupied with their cognitive abilities and programs have been developed to train these abilities (Berry et al., 2010; Fisher et al., 2010; Kesler et al., 2013; Klingberg et al., 2005). However, many training programs are impracticable since participants need to visit the laboratory or have to sit behind a personal computer. Mobile devices (e.g., smartphones, iPod touches), on the other hand, offer greater flexibility for self-supervised training, allowing the individual to train his/her cognitive abilities at any time or location (e.g., on the bus, in a café, at the hospital) and without direct supervision. Although, in recent years, mobile devices have found their way into psychological research, allowing for the collection of self-administered experimental data in real-life settings (e.g., Bless et al., 2013b; Dufau et al., 2011; Killingsworth and Gilbert, 2010, for a review,

see Miller, 2012), little is known about the applicability of these devices for training of cognitive functions.

The aim of the present study was to assess the feasibility of using a mobile application for self-supervised training of attention in healthy individuals. For this purpose, we have developed a mobile auditory attention training application for the iPhone/iPod touch, which is based on the forced-attention conditions of the consonant–vowel dichotic listening (CV-DL) paradigm (Hugdahl, 2003; Hugdahl and Andersson, 1986), and which has been validated in a previous study (Bless et al., 2013b). This paradigm requires participants to focus attention on syllables played in the right ear or left ear, while ignoring those played simultaneously in the other ear. In general, when confronted with two speech sounds at the same time, individuals tend to report the right-ear stimulus more often than the left-ear stimulus (e.g., Bryden, 1988; Hugdahl, 1995; Kimura, 1967), a phenomenon termed right-ear advantage (Shankweiler and Studdert-Kennedy, 1967). Thus, attending to and reporting the left-ear stimulus is considered to be more difficult than attending to and reporting the right-ear stimulus, posing different processing demands on auditory attention and cognitive control (Hugdahl et al., 2009; Kompus et al., 2012). As such, the paradigm serves

* Corresponding author at: Dept. of Biological and Medical Psychology, University of Bergen, Jonas Lies vei 91, 5009 Bergen, Norway. Tel.: +47 55586281.
E-mail address: josef.bless@psybp.uib.no (J.J. Bless).

as an analogue to everyday life situations, in which one auditory event, e.g. the response of our conversation partner, competes with multiple auditory streams, such as verbal utterances of other speakers or environmental sounds (traditionally referred to as the cocktail-party phenomenon) (Cherry, 1953). The ability to master such situations, however, is prone to individual differences (Conway et al., 2001), since deficits in auditory attention abilities have been reported in aging (Hugdahl et al., 2001; Passow et al., 2012; Takio et al., 2009), and in various clinical groups, e.g. schizophrenia patients (Hugdahl et al., 2013), preterm-born adolescents (Bless et al., 2013a), and children with dyslexia (Facoetti et al., 2003). Thus, the results of the current study, although based on healthy individuals, may offer new solutions for administering cognitive training in these patient groups. For example, feasibility of home-based cognitive training in schizophrenia patients has recently been reported (Hegde et al., 2012; Ventura et al., 2013).

Measuring the success of cognitive training can be approached on different levels of analysis. Certainly, improvement on the trained task itself has to be achieved. This has been shown previously using the CV-DL task (Soveri et al., 2013). However, such improvement is substantiated by investigating the neural correlates of the training effects in order to reveal underlying mechanisms (see Nyberg et al., 2003). For the present task and for mobile-application based cognitive training in general, this has to our knowledge not been explored previously. At the latest since Donald Hebb's theory of learning (Hebb, 1949), which has found due support by animal models (e.g., Bliss and Lomo, 1973; Kandel and Schwartz, 1982), behavioral changes (learning) are linked to neural mechanisms in the brain, mainly reflected in synaptic growth and neuronal firing. However, in recent years, this picture has been enriched by the notion that "less could be more" on both the anatomical (synaptic pruning) and the functional level (neural activation). One way of directly addressing this question is by examining training effects with functional magnetic resonance imaging (fMRI). Indeed, studies have shown that cognitive training results in both activation increases (e.g., Dahlin et al., 2008, Exp.1) as well as decreases (e.g., Schneiders et al., 2011); sometimes the same study reports increases in some and decreases in other regions (for a review, see Buschkuehl et al., 2012; Klingberg, 2010).

Thus, in the present study we expect successful training to be reflected by, (a) performance increase on the trained task, as well as (b) corresponding changes (increases or decreases) in neural activation in brain areas associated with speech perception and attention. In addition, to explore possible transfer of the trained task to other attentional tasks, we also included two cognitive interference tasks, one in the visual and one in the auditory domain.

2. Methods

2.1. Ethics statement

The study was approved by the Regional Committee for Medical Ethics in Western Norway (REK-Vest). All subjects gave written informed consent before the experiment and received financial compensation for their participation.

2.2. Subjects

Twenty-eight healthy subjects were recruited through university mailing lists and flyers posted on student blackboards. Subjects were randomly assigned to either the non-training control group ($N = 15$) or the training group ($N = 13$) in consecutive order, while accounting for a balanced sex distribution in both groups. Initially, subjects passed a screening for right-handedness (Edinburgh Handedness Inventory), no hearing impairment (Hughson–Westlake audiometric screening test), and no history of psychiatric disorder or neurological disease (self-report). There were no significant differences between the groups

Table 1
Group characteristics.

	Control	Training
Sex	8 males, 7 females	6 males, 7 females
Age	23.3 (± 0.6)	23.9 (± 0.7)
Handedness score	0.95 (± 0.02)	0.99 (± 0.02)
Hearing threshold (dB)	RE: 4.7 (± 1.7)	RE: 7.9 (± 1.9)
	LE: 3.7 (± 1.8)	LE: 6.2 (± 1.9)
Hearing threshold asymmetry (RE–LE)	2.6 (± 0.7)	3.2 (± 0.6)

Notes: No significant differences between the group means were found (for details see text). Measures of dispersion (given in parenthesis) are provided as standard error. RE = right ear, LE = left ear.

with regard to hearing acuity/asymmetry and handedness (all t < 1.2, all $p > .23$). For group characteristics, see Table 1. The reason for only including right-handers is to reduce between-subjects variations in lateralization, since evidence suggests that right-handers have more left-lateralized brains, as indicated by a stronger right-ear advantage compared to left-handers (see Van der Haegen et al., 2013).

2.3. Procedure overview

Subjects in the training group performed the auditory attention training with the mobile application for a period of 21 days, while control subjects did not receive training for the same time period. In order to assess the effectiveness of training, identical assessments were conducted in both groups on the first and last day of training/waiting. See Table 2 for an overview.

2.4. Training material and procedure

The training was conducted on iPod touch devices (4th generation) equipped with an in-house developed application, programmed in Xcode 4.2 (Apple Inc., Cupertino, CA). The sound stimuli were delivered via standard iPod earphones, which deliver output quality comparable to professional Sennheiser headphones (HD 280) with regard to interaural intensity differences [mean right–left differences for speech-relevant frequencies of 250 Hz–2 kHz are 0.32 dB (Apple) and −0.12 dB (Sennheiser)] (see Bless et al., 2013b). This is important for dichotic listening (DL) experiments since interaural differences above 6 dB affect the size of the ear advantage (Hugdahl et al., 2008). Subjects were instructed to adjust the main sound level with respect to the ambient noise condition during self-administration. The training task was based on the forced-attention conditions of the standard CV-DL paradigm (Hugdahl, 2003; Hugdahl and Andersson, 1986). In this paradigm, the stimulus material consists of six CV-syllables /ba/, /da/, /ga/, /ka/, /ta/, and /pa/ presented in pairs of all possible combinations, with different syllables presented simultaneously via earphones, one to the subject's right, and one to the left ear. The duration of the syllables was between 400 and 500 ms with an inter-stimulus interval of 4000 ms in which the subject had time to respond. The syllables were spoken by a male, native Norwegian speaker with constant voice intonation and intensity. The instructions were presented on the iPod touch screen to focus attention on and report the syllable heard in the right ear (forced-right condition, FR) only, or left ear (forced-left condition, FL) only. Responses were recorded by pressing the corresponding "button" on the touch screen (see Fig. 1). Conditions were presented in six blocks (3 FR, 3 FL) with five trials per block and each trial consisted of a different CV pair (see above). Thus, one training session included 60 trials. Feedback in terms of percentage correct reports was displayed on the device's screen following each session.

The training period was 21 days, and subjects were instructed to train in two sessions per day, one in the morning and one in the evening. On each occasion, the subjects themselves could decide where they would perform the training. Each training session lasted approx. 6 min, yielding a total training time of ca. 12 min per day. Results were saved

Table 2
Procedure overview.

Time point	Event	Training group	Control group
Day 1	*Screening*: Edinburgh Handedness Inventory Hughson–Westlake hearing test	x	x
	Transfer tasks: Cognitive interference tasks (visual and auditory)	x	x
	fMRI assessment: CV–DL task (NF, FR, FL)	x	x
Days 1–21	*Self-supervised cognitive training with iPod touch*: CV–DL task (FR, FL)	x	0
Day 21	*Transfer tasks and fMRI assessment* (see Day 1)	x	x

Notes: x = yes, 0 = no; NF = non-forced, FR = forced-right, FL = forced-left.

on the iPod touch and later extracted for analysis. The entire training period included an average number of 40.7 (±2.5) sessions, which is 97% of the required sessions. Fig. 2 shows the training course over the period of 21 days averaged across all members of the training group.

2.5. Transfer tasks

Transfer of the training was assessed with two cognitive interference tests (visual and auditory; computer-based), which were chosen for their similarity to the DL forced-attention conditions paradigm, in that they measure the ability to inhibit responses to irrelevant aspects of a stimulus situation. The visual interference test was based on the Stroop color–word test (MacLeod, 1991) with incongruent (e.g., the word "red" written in blue ink) and neutral stimuli (non-words written in various colors, e.g., "ba" written in red). The instruction to the subject was to report the color of the words, while ignoring their meaning. For each condition, there were 52 trials with stimulus durations of 2.0 s and inter-stimulus-intervals of 3.0 s. The auditory transfer task consisted of words presented via headphones in a congruent (e.g., hearing the word "low" spoken in a low pitch), incongruent (e.g., hearing the word "high", spoken in a low pitch) or neutral (e.g., hearing "ba" spoken in a high pitch) manner. The instruction was to report whether the words were spoken with a high or low pitch, while ignoring their meaning. The auditory transfer task consisted of 20, 52 and 52 trials respectively, with stimulus duration of 1.0 s and inter-stimulus-interval of 2.0 s. For both the visual and auditory training tasks, the interference effect was calculated by subtracting the mean response time of the neutral trials from the mean response time of the incongruent trials. Both transfer tests were administered from a laptop PC with the use of the E-Prime (2.0 Psychology Software Tools, Inc.) stimulus presentation software.

2.6. fMRI task

The task selected for the fMRI evaluation was the same as the training task. However, it included a third condition (non-forced, NF) with the instruction to listen to and report the syllable that was heard best (no attention focus). This served as a control condition, while the two forced-attention conditions (FR, FL) served as both training and outcome measures. Behavioral data recorded during scanning was scored as the number of correct reports for the right and left ear, respectively, and for the three conditions.

2.7. fMRI acquisition and preprocessing

MR imaging was performed on a 3 T GE Signa HDx scanner at Haukeland University Hospital in Bergen. Functional images were acquired using a T2-weighted gradient echo-planar imaging (EPI) sequence (TE = 30 ms; 90° flip angle) and were oriented to the structural image. A sparse sampling protocol was used [repetition time (TR) = 4.5 s, acquisition time (TA) = 1.5 s] with a silent gap for oral response of 3.0 s. All EPI volumes effectively covered the whole brain with 25 axial slices (0.5 mm interslice gap, 5.0 mm slice thickness, FOV 220 × 220, 64 × 64 scan matrix) and a voxel size of 3.44 mm × 3.44 mm × 5.0 mm. A T1-weighted FSPGR sequence, with standard parameters, was applied for 3D anatomy image acquisition.
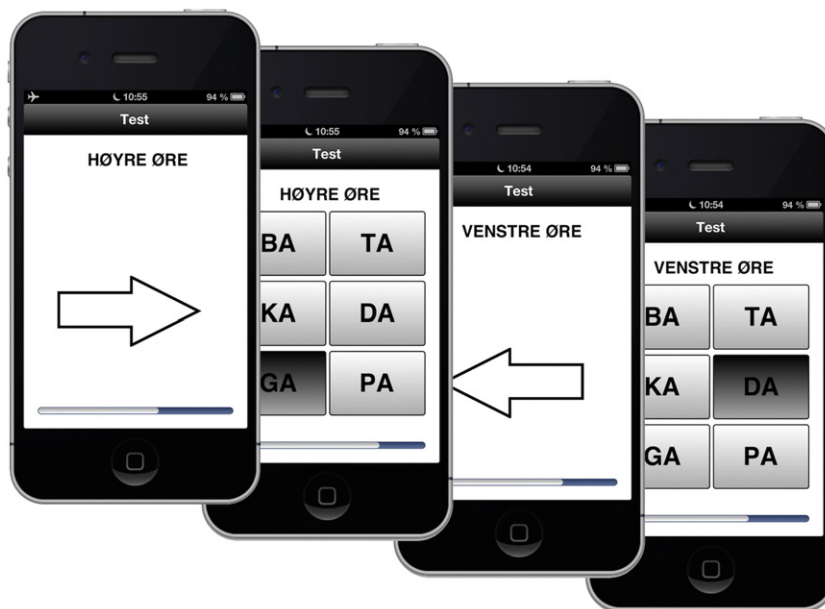

Fig. 1. Training application illustration. iPhone screenshots of the instruction-screens (focus on right ear, focus on left ear) and response-screens with answer buttons. The bar on the bottom of the screens depicts the remaining time for the current answer and until the next stimulus is presented.
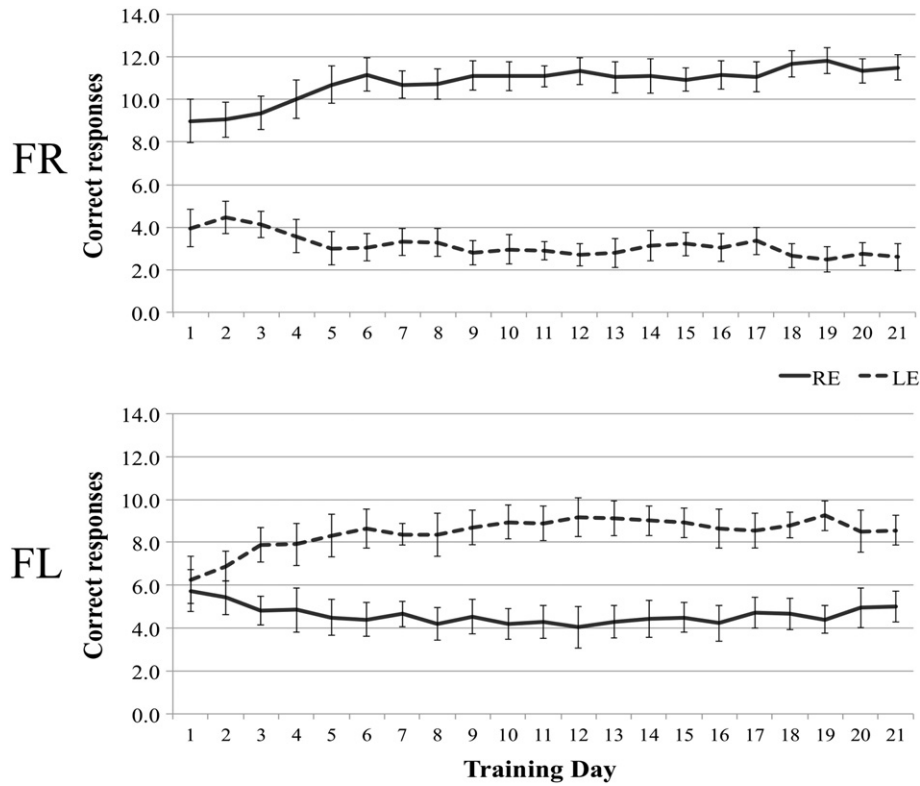
Fig. 2. Training progress — application results. Correct ear reports (y-axis) over the course of 21 training days (x-axis). Maximum correct reports = 15. FR = forced-right condition; FL = forced-left condition; RE = right ear; LE = left ear; error bars: standard error.

The experimental conditions were presented in blocks of 10 syllable-pairs and in two different orders, although always starting with the NF blocks, followed by alternating FR and FL blocks (order A: NF–NF–NF–FL–FR–FR–FL–FR–FL; order B: NF–NF–NF–FR–FL–FL–FR–FL–FR). The DL syllables were presented via MR compatible headphones (NordicNeuroLab Inc. http://nordicneurolab.com/) and subjects responded orally directly following stimulus presentation, still during the silent gap and before the next scan, thus avoiding interference of scanner noise and movement artifacts related to response articulation. The responses were recorded with an MR-compatible microphone and an MP3 player. The timing of the stimulus-presentations was controlled and synchronized with fMRI image acquisitions with the E-Prime software (Psychology Software Tools Inc., Pittsburgh, PA, USA).

The pre-processing of the fMRI data and all following analyses was performed with the Statistical Parametric Mapping (SPM 8) analysis software package (Wellcome Department of Cognitive Neurology, London, UK) running under MATLAB 2010b (Mathworks Inc., Natick, MA, USA). The EPI images were realigned to the first image in each time series and un-warping was performed to correct for the interaction of susceptibility artifacts and head movements. The unwarped mean images were then normalized to the MNI standard template and re-sampled to an isometric voxel size $2 \times 2 \times 2$ mm. Finally, using a 3D Gaussian filter of 8-mm FWHM, the normalized images were smoothed in order to compensate for the remaining inter-individual anatomical differences and increase the signal-to-noise-ratio.

For the analysis of activations for each individual, a SPM first-level analysis was done. This analysis was set up as a statistical model with one predictor for each of the three experimental conditions (NF, FR, FL). The predictors were convoluted with the canonical hemodynamic response function (hrf) and a temporal high pass filter (cutoff: 512 s) was applied. Individual movement parameters created during realignment were entered as multiple regressors into the first-level analysis as covariates of non-interest, in order to account for residual movement

artifacts. The estimated beta maps for the FR and FL predictors of each subject were submitted to second-level analysis (see next section).

2.8. Statistical analysis

2.8.1. Behavioral data analysis

In order to explore the effects of the auditory attention training on a behavioral level, the fMRI task results were analyzed using a $2 \times 3 \times 2 \times 2$ analysis of variance (ANOVA) with repeated-measures factors *Time point* (2 levels: pre- and post-training), *Condition* (3 levels: NF, FR, FL) and *Ear* (2 levels: RE, LE), as well as a between-subjects factor *Group* (training, control). The factorial ANOVA was followed up with appropriate lower-level ANOVAs and t-tests to test for simple main- and interaction-effects. The effects of interest were the interaction effects that included the factors *Group* and *Time point* for the dependent variable of the number of correct reports. Effect-sizes were calculated as eta-squared (η^2) statistics.

Data from the transfer tasks were analyzed with a 2×2 factorial ANOVA with repeated-measures factors *Time point* (2 levels: pre, post), and a between-subjects factor *Group* (training, control). The effect of interest was a *Group* $\times$ *Time point* interaction. The behavioral data was analyzed in SPSS 20.0 (IBM Corp., New York, USA).

2.8.2. BOLD fMRI data analysis

In order to investigate the neural response to the training, functional imaging data was analyzed based on a 2×2 factorial design with the repeated-measure factor *Time point* (pre- and post-training) and a between-subjects factor *Group* (training, control). A family-wise error correction (FWE) was applied to obtain a corrected significant threshold of $\alpha = 0.05$ and with an extent threshold of k = 15 voxels (approximated based on expected voxels per cluster: 12.6). The effect of interest was a *Group* $\times$ *Time point* interaction in the forced-attention conditions (FR, FL). As a follow-up to the whole-brain voxel-wise analysis and in

order to explore and visualize the effect in the regions that showed significant interactions, we extracted the hemodynamic response of the activation peaks using the MarsBaR toolbox (Brett et al., 2002).

3. Results

3.1. Behavioral results

Analysis of the fMRI task revealed a significant 4-way interaction of *Time point* (pre/post) × *Group* (control, training) × *Condition* (NF, FR, FL) × *Ear* (RE, LE) [$F(2,52) = 8.02$, $p = .001$, $\eta^2 = .01$]. When the two groups were analyzed separately in two 3-way ANOVAs as follow-up of the significant 4-way ANOVA, the training group showed a significant 3-way interaction between *Time point* (pre/post) × *Condition* (NF, FR, FL) × *Ear* (RE, LE) [$F(2,24) = 13.21$, $p < .001$, $\eta^2 = .06$], while no significant 3-way interaction was found in the control group [$F(2,28) = 0.57$, $p = .572$, $\eta^2 < .01$]. Thus, only the training group was further analyzed which showed significant effects for the 2-way interactions of *Time point* (pre/post) × *Ear* (RE, LE) in the FR condition [$F(1,12) = 14.34$, $p = .003$, $\eta^2 = .04$], and in the FL condition

[$F(1,12) = 12.58$, $p = .004$, $\eta^2 = .22$], respectively. Post-hoc paired t-tests of the pre–post correct ear report comparison showed significant differences for both the FR and FL conditions. This showed that the FR right-ear score was increased after training [$t(12) = 4.03$, $p < .01$, $d = 1.01$], as was the FL left-ear score [$t(12) = 3.51$, $p < .01$, $d = 1.23$], while the performance for the unattended ear was correspondingly suppressed [FR left-ear score: $t(12) = -2.67$, $p = .02$, $d = -0.85$, and FL right-ear score: $t(12) = -2.52$, $p = .03$, $d = -0.68$, respectively]. Fig. 3 shows the changes of ear scores from pre- to post-training in both groups.

Analysis of the transfer tasks showed no significant *Time point* × *Group* interaction effects, neither in the auditory [$F(1,26) = 0.80$, $p = .380$, $\eta^2 < .01$], nor in the visual interference test [$F(1,26) = 0.94$, $p = .341$, $\eta^2 = .01$].

3.2. BOLD fMRI results

The whole-brain fMRI analysis revealed significant effects of interest, i.e. interactions between *Group* × *Time point*, in several clusters corresponding to the forced-attention DL task (see Table 3, Fig. 4a). The FR
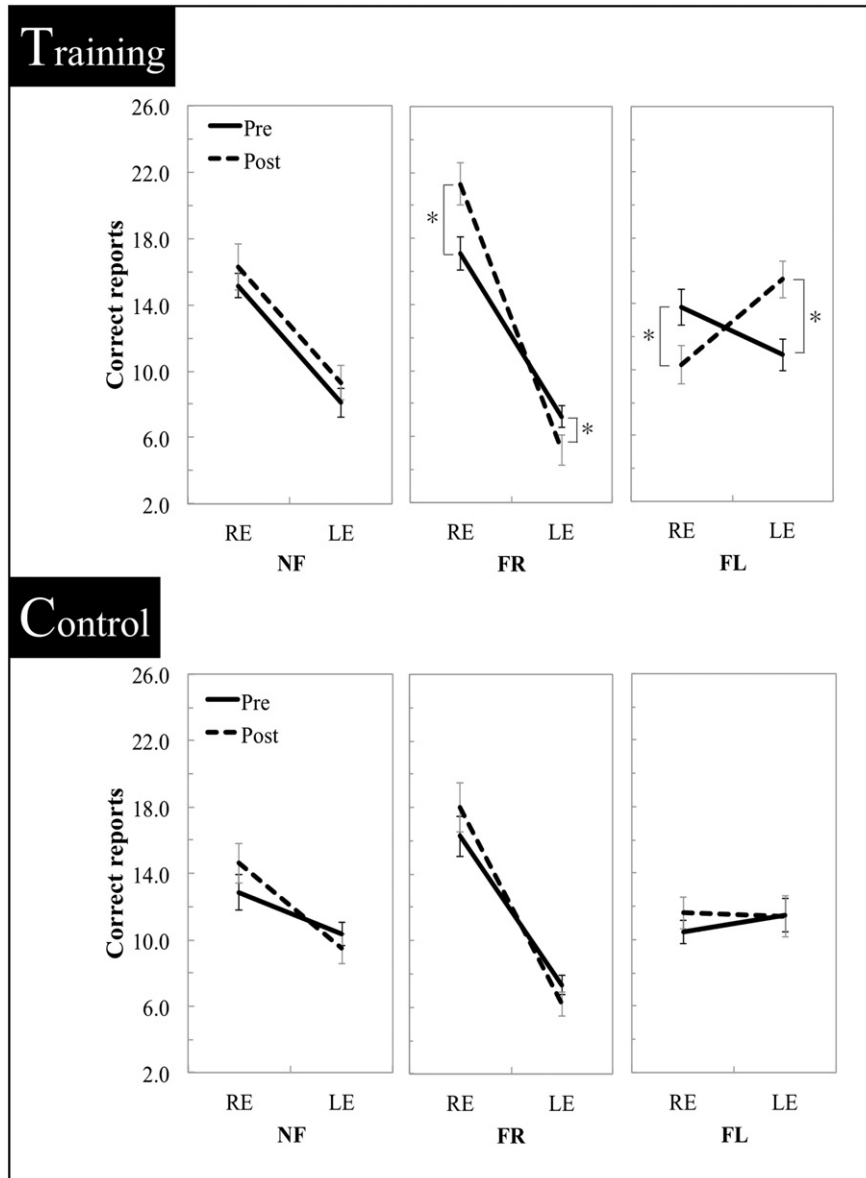

Fig. 3. fMRI task results. C = control group; T = training group. Correct ear reports measured pre- and post-training. Max. correct reports = 30. NF = non-forced; FR = forced-right; FL = forced-left; RE = right ear; LE = left ear; error bars: standard error. *$p < .05$.

Table 3
Peak location and anatomical labels of activation clusters for Group × Time point interactions in FR and FL conditions.

Cluster	Size (voxels)	t-Value	MNI coordinates			Anatomical label (for peak location)	Cluster extension
			x	y	z		
FR:							
1a	127	5.71	−52	−52	−6	Left ITG	Posterior regions
1b		4.27	−50	−64	−2	Left MTG	
2	45	5.61	−26	−74	−4	Left FG	Extends into the left LG
3	16	5.28	−34	−8	60	Left PG	
FL:							
1	26	5.15	−52	−52	−4	Left ITG	Bordering the left MTG
2	17	4.92	48	44	8	Right MFG	Stretches into the right IFG

Notes: ITG = inferior temporal gyrus, MTG = middle temporal gyrus, FG = fusiform gyrus, LG = lingual gyrus, PG = precentral gyrus, MFG = middle frontal gyrus, IFG = inferior frontal gyrus, FWE corrected ($p < .05$), extent threshold: k = 15 voxels, MNI = Montreal Neurol. Inst.

condition showed significant left-hemisphere interactions in the posterior inferior temporal gyrus (ITG) extending into the middle temporal gyrus (MTG), in the fusiform gyrus (FG) extending into the lingual gyrus (LG), and an additional cluster in the precentral gyrus (PG). The FL condition showed two significant interactions, one in the left posterior ITG bordering the MTG and another in the right middle frontal gyrus (MFG). The follow-up region-of-interest analysis showed that, in all clusters, the activation was decreased in the training group after training and increased in the control group (see Fig. 4b).

4. Discussion

The present findings showed that self-supervised training of auditory attention with a mobile application is both feasible and successful.

An improvement in attention performance was accompanied by corresponding change in brain activation, thus revealing a neural correspondence to the behavioral training effect. However, transfer of the training to other tasks tapping into similar mechanisms was not observed.

Beneficial effects of the training were reflected in the behavioral data at the end of the 21-day period, where the training group showed better performance in both forced-attention DL conditions compared to the control group. At the same time, no changes in the untrained, NF condition were observed, suggesting that training did not affect the right-ear advantage in the NF control condition. The forced-attention instruction may have an effect on two different stages of stimulus processing (Westerhausen et al., 2013). It may "proactively" interact with early auditory processing, e.g., by selectively attenuating the weight of the not-to-be-attended auditory input during encoding into working
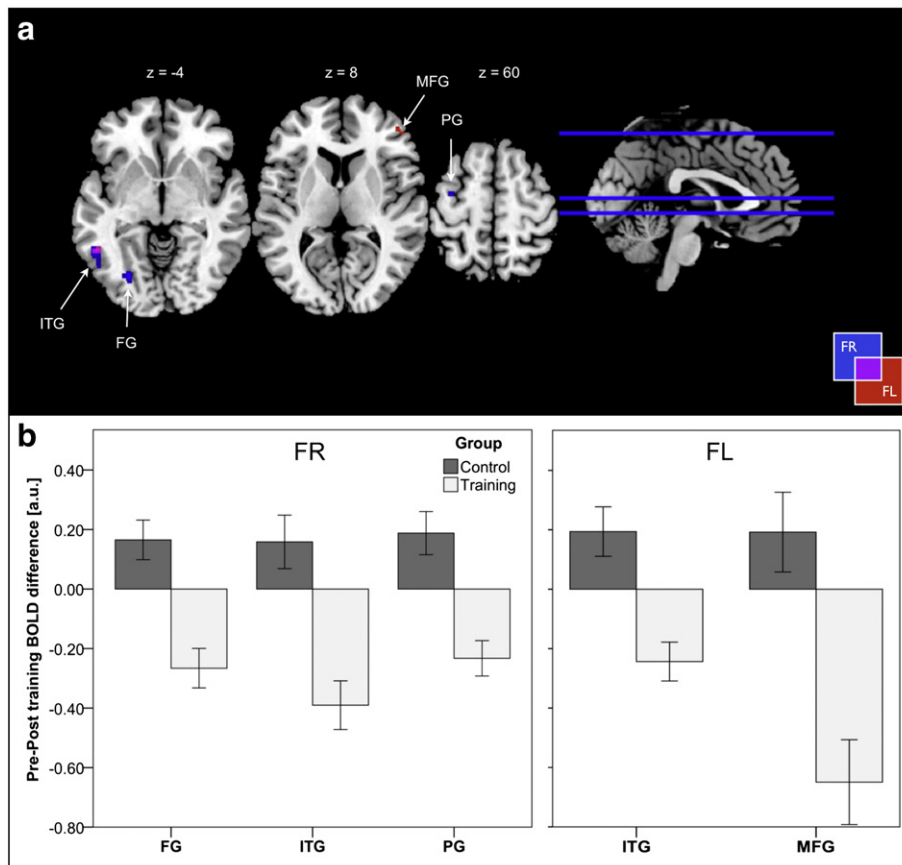

Fig. 4. Training effects as shown with BOLD fMRI. a) Brain regions showing significant Time point × Group interactions during the selective attention task. Color scheme: blue = forced-right condition, red = forced-left condition, purple = overlap of forced-right and forced-left condition. Activation maps were thresholded at $p < .05$, FWE corrected. Cluster extent threshold was set to k = 15 voxels. b) Activation changes from pre- to post-training for all significant clusters. Groups and conditions are displayed separately. FR = forced-right; FL = forced-left. a.u. = arbitrary units. FG = fusiform gyrus, ITG = inferior temporal gyrus, PG = precentral gyrus, MFG = middle frontal gyrus. Error bars: standard error.

memory (e.g., Alho et al., 1999; Treisman, 1964). However, the paradigm also relies on "reactive" executive control processes, necessary to resolve the competition between the two stimuli in working memory and follow the instruction to report the stimulus presented to a specific ear (Hiscock et al., 1999; Hugdahl et al., 2009). Accordingly, the observed training effects could have come from either or both stages, improving early attention processes and/or later executive control operations during stimulus selection. Interestingly, the effect-sizes indicate that training was particularly effective when attention was directed to the left ear. This may be the result of a ceiling effect in the less demanding FR condition, since the focus is on the dominant right-ear stimulus, already showing a high number of correct reports before the training. In the more effortful FL condition (Hugdahl et al., 2009), there is more room for improvement with the focus being on the "weaker" left-ear stimulus. This makes the FL task more susceptible for cognitive deficits (for a review, see Westerhausen and Hugdahl, 2010), yet also more receptive for training effects (see Soveri et al., 2013).

Regarding the fMRI findings, the training group displayed a post-training decrease in activation in regions associated with sensory and cognitive processing. Decreased activation after training is in line with the neural efficiency theory (Haier et al., 1992), interpreted to reflect more efficient use of cognitive resources as an effect of training, with fewer neurons and/or brain circuits necessary to meet the task demands. Training-induced decrease in activation in some areas has been reported previously (Erickson et al., 2007; Schneiders et al., 2011), and has also been observed in combination with increases in other areas (Dahlin et al., 2008; Olesen et al., 2004). The training-induced activation decreases may reflect more efficient processing at the neural level while the increases observed in the control group may be related to re-test effects. Thus it could be speculated that both groups followed an inverted U-shaped course with activation decreases after three weeks of training (training group) and activation increases as a result of repeated exposure to the task (control group) (see Hempel et al., 2004).

Relating the present training effects to previous studies using the same CV–DL paradigm in a conventional (non-training) fMRI experiment, substantial spatial overlap was found. The main cluster, located in the ITG/MTG region and modulated by training in both forced-attention conditions, has previously been associated with selective attention in DL (Hugdahl et al., 2000; Pugh et al., 1996). This regions also receives input from the planum temporale (Griffiths and Warren, 2002), an important region for phonetic processing (e.g., Binder et al., 1996; Jäncke et al., 2002; Uppenkamp et al., 2006). The FL condition revealed an additional significant cluster in the right MFG that stretched into the IFG. Generally, this region of the brain has been implicated in executive control functions such as inhibition (e.g., Aron et al., 2003; Rubia et al., 2003, for a review, see Verbruggen and Logan, 2008). Inhibition plays a role in the forced-attention conditions, since the signal from not-to-be-attended ear (e.g. the right ear during FL) needs to be suppressed. More specifically, the present training effects in the right frontal lobe is in line with previous CV–DL studies showing activation in the right IFG in response to instruction to focus on the left ear (Hugdahl et al., 2000; Jäncke and Shah, 2002; Thomsen et al., 2004). The observation that the frontal region only emerges in the FL supports the notion that processing of the FL condition taps into executive resources to a greater extent than the FR instruction (Hugdahl et al., 2009; Westerhausen and Hugdahl, 2010). Applied to the two-stage model (see above, Westerhausen et al., 2013), these findings suggest that training affects both early (temporal) and late (frontal) stages of auditory attention (see also, Ross et al., 2010; Larson and Lee, 2013).

The activation in the FG/LG was less intuitive given that the task is in the auditory domain. However, occipital lobe activations during auditory attention tasks have been reported previously using CVs (Kompus et al., 2012; Westerhausen et al., 2010) and tones (Cate et al., 2009). One explanation may be that these activations are the result of eye movements caused by the expectancy of a visual stimulus on the side of the attended ear. This might also explain the cluster in the PG and part of the primary motor cortex. It is unclear, however, why the effect is only seen in the FR condition.

The present results contribute to the growing body of research demonstrating the feasibility and effectiveness of mobile devices as platforms for self-administered behavioral intervention, previously assessed in a range of clinical contexts, e.g. management of alcohol abuse (Dulin et al., 2014), smoking cessation (Valdivieso-López et al., 2013), stress and mood problems (Lappalainen et al., 2013), and anxiety (Lindner et al., 2013; Pramana et al., 2014). Since interventions should be independent of time and place, and be available to individuals as they go about their daily lives, mobile devices appear to be the ideal instrument for delivery of these (Dagöö et al., 2014; Heron and Smyth, 2010, for a review, see Trull and Ebner-Priemer, 2013). In addition, mobile devices may be used to monitor the fluctuations of clinical symptoms, with a higher temporal resolution than the standard clinical interview. First attempts have been made to validate this method for assessment of mood and behavioral changes in outpatients with schizophrenia (Granholm et al., 2008; Palmier-Claus et al., 2012).

4.1. Limitations

A limitation to the use of mobile applications for cognitive training is the lack of control over the environment in which the training takes place, making results susceptible to unknown and potentially confounding factors. Nevertheless, it has recently been shown that results collected with a similar paradigm in real-life settings are comparable to those obtained in the laboratory (e.g., Bless et al., 2013b). For more control, one could use smartphone sensors to record contextual variables such as location, movement and ambient noise, and use these as covariates in the analysis, however, one should also be aware of the ethical implications of recording such data. Another challenge is the lack of control over the subject's training schedule. In order to minimize data omissions, it would be possible to schedule daily alarms on the training device. However, this would restrain the subjects' freedom to train wherever/whenever, the main advantage of using mobile devices in preference to stationary solutions.

Another question concerns the duration of the training. The observed performance plateau after 10 days suggests that a shorter period of training may be sufficient to induce plasticity effects. This could be addressed in future studies by including an intermediate fMRI assessment. Alternatively, to avoid ceiling effects, the task may include adjustments of difficulty level. For dichotic stimuli, this could be achieved by presenting the left and right ear stimuli with different intensities (Hugdahl et al., 2008; Westerhausen et al., 2009).

Furthermore, transfer to the untrained interference tasks was not observed despite the similarity of the tasks, i.e. both trained and transfer tasks tap into the inhibitory component of executive control, as outlined by Miyake et al. (2000). It may be speculated that the current sample size was too small: however, the lack of transfer may also be related to the specificity of the training task. According to the reverse hierarchy model proposed by Ahissar and Hochstein (2004), the degree of transfer or specificity indicates the level on which learning has occurred, i.e. training-induced neural changes on late stages of task processing yield clearer transfer effects. Although the present fMRI results showed that training affected both early and late stages of processing, early stages of auditory attention appeared to be affected in both conditions (ITG), suggesting that the training paradigm resulted in more specific and less general learning. This leads into an open discussion about what tasks are more likely to generalize to others and what factors facilitate transfer and why (see Green and Bavelier, 2008; Jaeggi et al., 2013; Melby-Lervåg and Hulme, 2013; Shipstead et al., 2012), which lies beyond the scope of this study. It should be noted, that, using a similar training paradigm, Soveri et al. (2013) found a transfer effect to an untrained auditory spatial attention task (in terms of decreased response errors).

5. Conclusions

The present findings suggest that mobile devices are feasible platforms for self-supervised cognitive training. The ability to focus on auditory input during competing stimulus conditions was increased as a result of training and accompanied by decreases in neural activation indicating more efficient stimulus processing at early and late stages of selective auditory attention. The greater flexibility that mobile-application based cognitive training offers over laboratory or PC-based training may be particularly beneficial in clinical settings where patients often have to follow strict routines and may not be able to leave the hospital facilities. This should encourage future research into the use of mobile applications for cognitive training with an emphasis on aging and various clinical populations with prominent auditory attention deficits.

Acknowledgments

References

Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. 8, 457–464.

Alho, K., Medvedev, S.V., Pakhomov, S.V., Roudas, M.S., Tervaniemi, M., Reinikainen, K., Zeffiro, T., Näätänen, R., 1999. Selective tuning of the left and right auditory cortices during spatially directed attention. Brain Res. Cogn. Brain Res. 7, 335–341.

Aron, A.R., Fletcher, P.C., Bullmore, E.T., Sahakian, B.J., Robbins, T.W., 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat. Neurosci. 6, 115–116.

Berry, A.S., Zanto, T.P., Clapp, W.C., Hardy, J.L., Delahunt, P.B., Mahncke, H.W., Gazzaley, A., 2010. The influence of perceptual training on working memory in older adults. PLoS One 5, e11537.

Binder, J.R., Frost, J.A., Hammeke, T.A., Rao, S.M., Cox, R.W., 1996. Function of the left planum temporale in auditory and linguistic processing. Brain 119 (Pt 4), 1239–1247.

Bless, J.J., Hugdahl, K., Westerhausen, R., Løhaugen, G.C., Eidheim, O.C., Brubakk, A.M., Skranes, J., Gramstad, A., Håberg, A.K., 2013a. Cognitive control deficits in adolescents born with very low birth weight (≤1500 g): evidence from dichotic listening. Scand. J. Psychol. 54, 179–187.

Bless, J.J., Westerhausen, R., Arciuli, J., Kompus, K., Gudmundsen, M., Hugdahl, K., 2013b. "Right on all occasions?" — On the feasibility of laterality research using a smartphone dichotic listening application. Front. Psychol. 4, 42.

Bliss, T.V., Lomo, T., 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. 232, 331–356.

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of interest analysis using an SPM toolbox. International Conference on Functional Mapping of the Human Brain. John Wiley & Sons, Inc., Senai.

Bryden, M.P., 1988. An overview of the dichotic listening procedure and its relation to cerebral organization. In: Hugdahl, K. (Ed.), Handbook of Dichotic Listening: Theory, Methods and Research. John Wiley & Sons, Oxford, England.

Buschkuehl, M., Jaeggi, S.M., Jonides, J., 2012. Neuronal effects following working memory training. Dev. Cogn. Neurosci. 2 (Suppl. 1), S167–S179.

Cate, A.D., Herron, T.J., Yund, E.W., Stecker, G.C., Rinne, T., Kang, X., Petkov, C.I., Disbrow, E.A., Woods, D.L., 2009. Auditory attention activates peripheral visual cortex. PLoS One 4, e4645.

Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with 2 ears. J. Acoust. Soc. Am. 25, 975–979.

Conway, A.A., Cowan, N., Bunting, M., 2001. The cocktail party phenomenon revisited: the importance of working memory capacity. Psychon. Bull. Rev. 8, 331–335.

Dagöö, J., Asplund, R.P., Bsenko, H.A., Hjerling, S., Holmberg, A., Westh, S., Öberg, L., Ljotsson, B., Carlbring, P., Furmark, T., Andersson, G., 2014. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. J. Anxiety Disord. 28, 410–417.

Dahlin, E., Neely, A.S., Larsson, A., Bäckman, L., Nyberg, L., 2008. Transfer of learning after updating training mediated by the striatum. Science 320, 1510–1512.

Dufau, S., Dunabeitia, J.A., Moret-Tatay, C., Mcgonigal, A., Peeters, D., Alario, F.X., Balota, D.A., Brysbaert, M., Carreiras, M., Ferrand, L., Ktori, M., Perea, M., Rastle, K., Sasburg, O., Yap, M.J., Ziegler, J.C., Grainger, J., 2011. Smart phone, smart science: how the use of smartphones can revolutionize research in cognitive science. PLoS One 6, e24974.

Dulin, P.L., Gonzalez, V.M., Campbell, K., 2014. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: usability and early outcomes. Subst. Abus. 35, 168–175.

Erickson, K.I., Colcombe, S.J., Wadhwa, R., Bherer, L., Peterson, M.S., Scalf, P.E., Kim, J.S., Alvarado, M., Kramer, A.F., 2007. Training-induced functional activation changes in dual-task processing: an FMRI study. Cereb. Cortex 17, 192–204.

Facoetti, A., Lorusso, M.L., Paganoni, P., Cattaneo, C., Galli, R., Umilta, C., Mascetti, G.G., 2003. Auditory and visual automatic attention deficits in developmental dyslexia. Brain Res. Cogn. Brain Res. 16, 185–191.

Fisher, M., Holland, C., Subramaniam, K., Vinogradov, S., 2010. Neuroplasticity-based cognitive training in schizophrenia: an interim report on the effects 6 months later. Schizophr. Bull. 36, 869–879.

Granholm, E., Loh, C., Swendsen, J., 2008. Feasibility and validity of computerized ecological momentary assessment in schizophrenia. Schizophr. Bull. 34, 507–514.

Green, C.S., Bavelier, D., 2008. Exercising your brain: a review of human brain plasticity and training-induced learning. Psychol. Aging 23, 692–701.

Griffiths, T.D., Warren, J.D., 2002. The planum temporale as a computational hub. Trends Neurosci. 25, 348–353.

Haier, R.J., Siegel, B., Tang, C., Abel, L., Buchsbaum, M.S., 1992. Intelligence and changes in regional cerebral glucose metabolic rate following learning. Intelligence 16, 415–426.

Hebb, D.O., 1949. The Organization of Behavior: A Neuropsychological Theory. Wiley, New York.

Hegde, S., Rao, S.L., Raguram, A., Gangadhar, B.N., 2012. Addition of home-based cognitive retraining to treatment as usual in first episode schizophrenia patients: a randomized controlled study. Indian J. Psychiatry 54, 15–22.

Hempel, A., Giesel, F.L., Garcia Caraballo, N.M., Amann, M., Meyer, H., Wüstenberg, T., Essig, M., Schröder, J., 2004. Plasticity of cortical activation related to working memory during training. Am. J. Psychiatry 161, 745–747.

Heron, K.E., Smyth, J.M., 2010. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br. J. Health Psychol. 15, 1–39.

Hiscock, M., Inch, R., Kinsbourne, M., 1999. Allocation of attention in dichotic listening: differential effects on the detection and localization of signals. Neuropsychology 13, 404–414.

Hugdahl, K., 1995. Dichotic Listening: Probing Temporal Lobe Functional Integrity. In: Davidson, R.J., Hugdahl, K. (Eds.), MIT Press, Cambridge, Mass.

Hugdahl, K., 2003. Dichotic listening in the study of auditory laterality. In: Hugdahl, K., Davidson, R.J. (Eds.), The Asymmetrical Brain. MIT Press, Cambridge, Mass.

Hugdahl, K., Andersson, L., 1986. The "forced-attention paradigm" in dichotic listening to CV-syllables: a comparison between adults and children. Cortex 22, 417–432.

Hugdahl, K., Law, I., Kyllingsbaek, S., Brønnick, K., Gade, A., Paulson, O.B., 2000. Effects of attention on dichotic listening: an 15O-PET study. Hum. Brain Mapp. 10, 87–97.

Hugdahl, K., Carlsson, G., Eichele, T., 2001. Age effects in dichotic listening to consonant-vowel syllables: interactions with attention. Dev. Neuropsychol. 20, 445–457.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., Hämäläinen, H., 2008. The effect of stimulus intensity on the right ear advantage in dichotic listening. Neurosci. Lett. 431, 90–94.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., Laine, M., Hämäläinen, H., 2009. Attention and cognitive control: unfolding the dichotic listening story. Scand. J. Psychol. 50, 11–22.

Hugdahl, K., Nygård, M., Falkenberg, L.E., Kompus, K., Westerhausen, R., Kroken, R., Johnsen, E., Løberg, E.M., 2013. Failure of attention focus and cognitive control in schizophrenia patients with auditory verbal hallucinations: evidence from dichotic listening. Schizophr. Res. 147, 301–309.

Jaeggi, S.M., Buschkuehl, M., Shah, P., Jonides, J., 2013. The role of individual differences in cognitive training and transfer. Mem. Cogn. 42, 464–480.

Jäncke, L., Shah, N.J., 2002. Does dichotic listening probe temporal lobe functions? Neurology 58, 736–743.

Jäncke, L., Wüstenberg, T., Scheich, H., Heinze, H.J., 2002. Phonetic perception and the temporal cortex. Neuroimage 15, 733–746.

Kandel, E.R., Schwartz, J.H., 1982. Molecular biology of learning: modulation of transmitter release. Science 218, 433–443.

Kesler, S., Hadi Hosseini, S.M., Heckler, C., Janelsins, M., Palesh, O., Mustian, K., Morrow, G., 2013. Cognitive training for improving executive function in chemotherapy-treated breast cancer survivors. Clin. Breast Cancer 13, 299–306.

Killingsworth, M.A., Gilbert, D.T., 2010. A wandering mind is an unhappy mind. Science 330, 932.

Kimura, D., 1967. Functional asymmetry of the brain in dichotic listening. Cortex 3, 163–178.

Klingberg, T., 2010. Training and plasticity of working memory. Trends Cogn. Sci. 14, 317–324.

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C.G., Forssberg, H., Westerberg, H., 2005. Computerized training of working memory in children with ADHD — a randomized, controlled trial. J. Am. Acad. Child Adolesc. Psychiatry 44, 177–186.

Kompus, K., Specht, K., Ersland, L., Juvodden, H.T., Van Wageningen, H., Hugdahl, K., Westerhausen, R., 2012. A forced-attention dichotic listening fMRI study on 113 subjects. Brain Lang. 121, 240–247.

Lappalainen, P., Kaipainen, K., Lappalainen, R., Hoffrén, H., Myllymäki, T., Kinnunen, M.L., Mattila, E., Happonen, A.P., Rusko, H., Korhonen, I., 2013. Feasibility of a personal health technology-based psychological intervention for men with stress and mood problems: randomized controlled pilot trial. JMIR Res. Protocol. 2, e1.

Larson, E., Lee, A.K., 2013. The cortical dynamics underlying effective switching of auditory spatial attention. Neuroimage 64, 365–370.

Lindner, P., Ivanova, E., Ly, K.H., Andersson, G., Carlbring, P., 2013. Guided and unguided CBT for social anxiety disorder and/or panic disorder via the Internet and a smartphone application: study protocol for a randomised controlled trial. Trials 14, 437.

Macleod, C.M., 1991. Half a century of research on the Stroop effect: an integrative review. Psychol. Bull. 109, 163–203.

Melby-Lervåg, M., Hulme, C., 2013. Is working memory training effective? A meta-analytic review. Dev. Psychol. 49, 270–291.

Miller, G., 2012. The smartphone psychology manifesto. Perspect. Psychol. Sci. 7, 221–237.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cogn. Psychol. 41, 49–100.

Nyberg, L., Sandblom, J., Jones, S., Neely, A.S., Petersson, K.M., Ingvar, M., Bäckman, L., 2003. Neural correlates of training-related memory improvement in adulthood and aging. Proc. Natl. Acad. Sci. U. S. A. 100, 13728–13733.

Olesen, P.J., Westerberg, H., Klingberg, T., 2004. Increased prefrontal and parietal activity after training of working memory. Nat. Neurosci. 7, 75–79.

Palmier-Claus, J.E., Ainsworth, J., Machin, M., Barrowclough, C., Dunn, G., Barkus, E., Rogers, A., Wykes, T., Kapur, S., Buchan, I., Salter, E., Lewis, S.W., 2012. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. BMC Psychiatr. 12, 172.

Passow, S., Westerhausen, R., Wartenburger, I., Hugdahl, K., Heekeren, H.R., Lindenberger, U., Li, S.C., 2012. Human aging compromises attentional control of auditory perception. Psychol. Aging 27, 99–105.

Pramana, G., Parmanto, B., Kendall, P.C., Silk, J.S., 2014. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed. J. E Health 20, 419–427.

Pugh, K.R., Shaywitz, B.A., Shaywitz, S.E., Fulbright, R.K., Byrd, D., Skudlarski, P., Shankweiler, D.P., Katz, L., Constable, R.T., Fletcher, J., Lacadie, C., Marchione, K., Gore, J.C., 1996. Auditory selective attention: an fMRI investigation. Neuroimage 4, 159–173.

Ross, B., Hillyard, S.A., Picton, T.W., 2010. Temporal dynamics of selective attention during dichotic listening. Cereb. Cortex 20, 1360–1371.

Rubia, K., Smith, A.B., Brammer, M.J., Taylor, E., 2003. Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. Neuroimage 20, 351–358.

Schneiders, J.A., Opitz, B., Krick, C.M., Mecklinger, A., 2011. Separating intra-modal and across-modal training effects in visual working memory: an fMRI investigation. Cereb. Cortex 21, 2555–2564.

Shankweiler, D., Studdert-Kennedy, M., 1967. Identification of consonants and vowels presented to left and right ears. Q. J. Exp. Psychol. 19 (59-&).

Shipstead, Z., Redick, T.S., Engle, R.W., 2012. Is working memory training effective? Psychol. Bull. 138, 628–654.

Soveri, A., Tallus, J., Laine, M., Nyberg, L., Bäckman, L., Hugdahl, K., Tuomainen, J., Westerhausen, R., Hämäläinen, H., 2013. Modulation of auditory attention by training: evidence from dichotic listening. Exp. Psychol. 60, 44–52.

Takio, F., Koivisto, M., Jokiranta, L., Rashid, F., Kallio, J., Tuominen, T., Laukka, S.J., Hämäläinen, H., 2009. The effect of age on attentional modulation in dichotic listening. Dev. Neuropsychol. 34, 225–239.

Thomsen, T., Rimol, L.M., Ersland, L., Hugdahl, K., 2004. Dichotic listening reveals functional specificity in prefrontal cortex: an fMRI study. Neuroimage 21, 211–218.

Treisman, A.M., 1964. Verbal cues, language, and meaning in selective attention. Am. J. Psychol. 77, 206–219.

Trull, T.J., Ebner-Priemer, U., 2013. Ambulatory assessment. Annu. Rev. Clin. Psychol. 9, 151–176.

Uppenkamp, S., Johnsrude, I.S., Norris, D., Marslen-Wilson, W., Patterson, R.D., 2006. Locating the initial stages of speech-sound processing in human temporal cortex. Neuroimage 31, 1284–1296.

Valdivieso-López, E., Flores-Mateo, G., Molina-Gómez, J.D., Rey-Renones, C., Barrera Uriarte, M.L., Duch, J., Valverde, A., 2013. Efficacy of a mobile application for smoking cessation in young people: study protocol for a clustered, randomized trial. BMC Public Health 13, 704.

Van Der Haegen, L., Westerhausen, R., Hugdahl, K., Brysbaert, M., 2013. Speech dominance is a better predictor of functional brain asymmetry than handedness: a combined fMRI word generation and behavioral dichotic listening study. Neuropsychologia 51, 91–97.

Ventura, J., Wilson, S.A., Wood, R.C., Hellemann, G.S., 2013. Cognitive training at home in schizophrenia is feasible. Schizophr. Res. 143, 397–398.

Verbruggen, F., Logan, G.D., 2008. Response inhibition in the stop-signal paradigm. Trends Cogn. Sci. 12, 418–424.

Westerhausen, R., Hugdahl, K., 2010. Cognitive control of auditory laterality. In: Hugdahl, K., Westerhausen, R. (Eds.), The Two Halves of the Brain: Information Processing in the Cerebral Hemispheres. MIT Press, Cambridge, Mass.

Westerhausen, R., Moosmann, M., Alho, K., Medvedev, S., Hämäläinen, H., Hugdahl, K., 2009. Top-down and bottom-up interaction: manipulating the dichotic listening ear advantage. Brain Res. 1250, 183–189.

Westerhausen, R., Moosmann, M., Alho, K., Belsby, S.O., Hämäläinen, H., Medvedev, S., Specht, K., Hugdahl, K., 2010. Identification of attention and cognitive control networks in a parametric auditory fMRI study. Neuropsychologia 48, 2075–2081.

Westerhausen, R., Passow, S., Kompus, K., 2013. Reactive cognitive-control processes in free-report consonant–vowel dichotic listening. Brain Cogn. 83, 288–296.