Within-subject consistency of paired associative stimulation as assessed by linear mixed models.

Myrthe Julia Ottenhoff[a,b], Lana Fani[a,b], Nicole Stephanie Erler[c], Jesminne Castricum[a,b,d], Imara Fedora Obdam[d], Thijs van der Vaart[a,b,e], Steven Aaron Kushner[b,d], Marie-Claire Yvette de Wit[b,e], Ype Elgersma[b,f*#], J.H.M. Tulen[b,g*#]

[a] Department of Neuroscience, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[b] ENCORE Expertise Center for Neurodevelopmental Disorders, Erasmus MC – Sophia Children's Hospital, University Medical Center Rotterdam, Rotterdam, The Netherlands

[c] Department of Biostatistics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[d] Department of Psychiatry, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[e] Department of Neurology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[f] Department of Neuroscience, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[g] Department of Psychiatry, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

* corresponding authors

E-mail: y.elgersma@erasmusmc.nl; j.h.m.tulen@erasmusmc.nl

[#] YE and JHMT are joint senior authors.

Abstract

Objective

Paired associative stimulation (PAS) is a TMS paradigm used to induce long-term potentiation in the human cortex. Little is known about the within-subject consistency of PAS-induced effects. We determined PAS-induced effects and their consistency in healthy volunteers between two PAS sessions. Additionally, we assessed the benefit of applying linear mixed models (LMMs) to PAS data.

Methods

Thirty-eight healthy volunteers underwent two identical PAS sessions with a >1 week interval. During each session, motor evoked potentials (MEPs) were assessed once before PAS induction and 3 times after at 30 min intervals.

Results

We did not detect any significant potentiation of MEP size after PAS induction. However, MEP size during PAS induction showed significant potentiation over time in both sessions ($LR(1) = 13.36$, $p<0.001$). Nevertheless, there was poor within-subject consistency of PAS-induced effects both during (ICC=0.15) and after induction (ICC=0.03-0.08). Additionally, statistical model selection procedures demonstrate that LMMs are more appropriate than conventional longitudinal models for estimating PAS-induced effects ($LR(34) = 214.73$, $p < 0.001$).

Conclusion

PAS-induced effects are more pronounced during than after induction, have a low within-subject consistency in any phase of the measurement, and are best estimated with LMMs. The implication of our study is that PAS is an inappropriate method to assess the evolution of brain plasticity over time periods longer than the PAS measurement itself.

Introduction

Synaptic plasticity is a fundamental process in our central nervous system, because it is essential for learning and memory [1, 2]. In addition, plasticity deficits are important in the etiology of neurocognitive disorders [3, 4]. Synaptic plasticity is conventionally measured with invasive techniques, such as field electrophysiology, which cannot readily be performed in human subjects. Non-invasive extracranial brain stimulation techniques, such as transcranial magnetic stimulation (TMS), have emerged in the last two decades, enabling measuring plasticity-like effects in human subjects. Specific TMS paradigms have been shown to produce long-term potentiation (LTP) as well as long-term depression (LTD) like effects, resembling the plasticity effects seen using intraparenchymal electrophysiological studies in animal models.

Paired associative stimulation (PAS) is one of the TMS paradigms used to induce long-term plasticity [5]. PAS is typically applied by pairing median nerve stimulation (MNS) with magnetic stimulation of the contralateral hand area of the primary motor cortex (M1). Consistent with the fundamental properties of spike-timing dependent plasticity (STDP) [7], when MNS precedes magnetic stimulations by 25ms, PAS stimulation induces a long-term increase in excitability of the M1 hand area that can be observed as an increase of motor-evoked potentials (MEPs) in the contralateral hand. In contrast, if the MNS precedes the magnetic stimulation by 10ms, the result is a long-term depression effect [6]. The resemblance to STDP is further strengthened by pharmacological studies showing that PAS is dependent on the function of the *N*-methyl-D-aspartate (NMDA) receptor, known to be essential for many forms of long-term synaptic plasticity [7].

Because of the temporal similarity of PAS results to classical STDP experiments in rodents, PAS has emerged as a potentially very useful proxy for studying long-term synaptic plasticity in human subjects. However, PAS produces highly variable results in humans [8, 9], which is often attributed to the challenge of achieving similar levels of standardization as for animal experiments. Environmental factors, lifestyle factors, experimental conditions and even genetic determinants have been suggested

to influence the magnitude of the PAS-induced plasticity (for review see Ridding et al. (2010) and Wischnewski et al. (2016)) [10, 11]. However, these factors only explain between-subject variability, whereas little is known about the within-subject consistency.

Besides inter- and intra-individual variability, PAS studies show variable effect sizes between laboratories as well [9, 11]. In addition to optimizing experimental procedures, some types of variability might be possible to account for by appropriate statistical modeling. PAS measurements generate relatively complex data, combining both repeated measures as well as a nested data structure (i.e. multiple MEP size assessments per time point). In the last decades, linear mixed models (LMMs) have emerged as a statistical method that is specifically suited to handle such a data structure, reducing the chance of both false-positive and false-negative results [12, 13]. Additionally, LMMs are excellent for estimating reproducibility measures in the form of intra-class correlations. However, the utility of LMMs for the analysis of PAS data has not yet been demonstrated.

In this study, we examined healthy volunteers using two identical PAS sessions with an interval of at least 1 week, in order to determine the within-subject consistency between the two sessions. Using linear mixed model (LMM) analyses, we estimated PAS-induced effects and PAS consistency.

Materials and Methods

Subjects

Thirty-eight out of 61 subjects were included in this study (reasons for exclusion are summarized in Table S1 Table), who were recruited by advertising in the local community and on a Dutch research subject-recruitment website. Subjects were included if aged 18-40, right-handed according to the Edinburgh Handedness Inventory [14], in good health, medication free (excluding contraceptives) and able and willing to give written informed consent. Subjects were excluded if they were women lactating or pregnant, had a history of psychiatric illness and/or treatment, had a history of

neurological illness or did not meet the international safety guidelines considering TMS [15, 16]. All subjects underwent the Wechsler Abbreviated Scale of Intelligence (WASI) [17] to determine their intelligence quotient (IQ) [18] for descriptive purposes. This study was approved by the Medical Ethical Review Board of the Erasmus MC Rotterdam in 2013, requiring study procedures to comply with the latest version of the Declaration of Helsinki.

Fig 1. Schematic representation of the PAS paradigm.

(A) Schematic of one PAS session, in which the PAS induction is preceded by a baseline measurement consisting of 20 TMS stimulations. To measure the change in MEP-size as of PAS induction, this measurement is repeated 3 times with 30 min intervals. (B) Experimental setup during PAS induction, where median nerve stimulation (MNS) at 3 times the sensory threshold (ST) precedes transcranial magnetic stimulations (TMS) delivered at the stimulation intensity 1mV (SI1mV) by 25 ms. Motor evoked potentials (MEPs) are measured using electromyography (EMG) of the abductis pollicis brevis muscle. **C** Example traces of single MEPs before and after PAS induction.

Electromyography

Muscle activity was recorded from the left abductor pollicis brevis (ABP) muscle with electromyography (EMG), using Ag-AgCl electrodes in a belly-tendon montage. EMG signals were amplified using a universal amplifier (ANT Neuro, Enschede, The Netherlands) and digitalized at 5kHz for later offline analysis using Visor2 XT software (ANT Neuro, Enschede, The Netherlands). During measurements, a continuous EMG signal and trigger related EMG epochs were plotted at real time for online analysis, while applying a 50Hz notch filter and a 20-2000Hz bandpass filter.

Transcranial magnetic stimulation

Subjects were invited in the afternoon between 12 and 5.30 PM [19], and were asked to not perform intense physical activities 24 hours prior to the measurement and to not smoke nicotine cigarettes or drink coffee on the day of the measurement. They were seated in a comfortable chair with their left

arm resting on a pillow and were told to maximally relax their left hand during the measurement. Magnetic stimulations were applied using a figure-of-eight coil with an inner diameter of 27mm and outer diameter of 97mm, connected to a MagPro X100 with MagOption TMS device (MagVenture, Farum, Denmark). The coil was held tangentially to the left primary cortex and diverging 45˚ from midline. The electric field subsequently created in the cortex had a posterior to anterior direction.

To find the optimal position of the coil in order to maximally activate the ABP (the hotspot), TMS stimulations were randomly placed around a predefined reference point. This reference point was defined as the location at 10% of the ear-to-ear span lateral to Cz over the right hemisphere. Data on coil location and position at every stimulation was collected using a neuronavigation system (ANT Neuro, Enschede, The Netherlands), allowing a precise spatial definition of the hotspot and precise determination of the angle and distance errors of every stimulation relative to the hotspot. All TMS procedures hereafter described are performed at the hotspot.

The resting motor threshold (RMT) was determined using a maximum-likelihood threshold hunting procedure [20]. For this procedure, a MEP with a peak-to-peak amplitude of ≥50μV with an onset within a timeframe of 20-50ms post stimulation was considered a MEP. Subsequently, the stimulation intensity 1mV (SI1mV) was determined, defined as the percentage of maximal stimulation output (%MSO) of the TMS device that resulted in a mean MEP of 0.8 - 1.2 mV after 10 stimulations at 0.1Hz.

Paired associative stimulation

Subjects underwent two identical paired associative stimulation (PAS) sessions at least one week apart. First, baseline cortical excitability was assessed by applying a train of 20 magnetic stimulations at the SI1mV at 0.1Hz. Second, the PAS induction phase followed, consisting of 200 paired stimulations at 0.25Hz with each pair consisting of electric MNS followed by magnetic stimulation with an inter stimulus interval of 25ms. Third, after this plasticity induction phase, the cortical

159 excitability measurement at baseline was repeated at three time points: immediately (Post 1), 30
160 minutes (Post 2), and 60 minutes (Post 3) after completing PAS induction (Fig 1B).
161
162 MNS during the PAS-induction was applied with a strength of three times the sensory threshold, using
163 a bipolar bar electrode connected to a constant current stimulator (model DS7A; Digitimer Ltd.,
164 Letchworth Garden City, UK) (Fig 1A). In case this stimulation intensity surpassed the pain threshold,
165 it was lowered to a painless but clearly noticeable level. The subject's attention level was standardized
166 as much as possible by applying four randomly timed electric stimuli during PAS induction to the
167 middle phalanx of the left thumb, and instructing participants upfront of PAS induction to focus their
168 attention on their left thumb, to count the number of thumb stimuli sensed and to report this number
169 after PAS induction [21]. These electric stimulations were administered at two times the sensory
170 threshold using a double ring electrode connected to a constant current stimulator (Micromed S.p.A,
171 Mogliano Veneto, Italy).
172

173 MEP-analysis

174 The EMG signal for every magnetic stimulation applied was stored for offline analysis as epochs of -
175 300ms to +300ms surrounding the TMS trigger (Fig 1C). Using software programmed in LabVIEW
176 (National Instruments, Austin, TX, US) pre-MEP noise, the maximal peak-to-peak amplitude and
177 MEP onset were determined using a seven-step data processing procedure:

178 1. Signals were linearly detrended.
179 2. The average amplitude value of the -300ms to -20ms before the TMS trigger was subtracted
180 to create a zero-baseline.
181 3. To prevent ringing due to filtering, the stimulation artefact was removed between -2ms to
182 +4ms surrounding the TMS trigger, followed by linear interpolation. For MEPs obtained
183 during PAS induction, the stimulation artefact of the MNS was removed similarly.
184 4. Signals were filtered using both a 20-2000Hz bandpass filter and a 50Hz-notch filter.

5. Noise was determined on a time window of -25ms to +15ms surrounding the TMS trigger that was depleted of baseline wandering by subtracting a 2nd-order polynomial fit. It was defined as a peak-to-peak amplitude of >50μV or an SD of >15. Signals meeting these criteria were discarded for further statistical analysis.
6. The maximal peak-to-peak amplitude of every MEP was determined within a 20-48ms time frame following the TMS trigger.
7. The MEP onset was determined within an 18-30ms time frame following the TMS trigger.

Statistical analysis

Statistical analyses were performed using R version 3.3.3 [22] and the nlme package for mixed model analysis specifically [23]. Session specific subject characteristics were compared using paired t-tests for normally distributed data (RMT and SI1mV), a Wilcoxon Signed Rank test for non-normal continuous data (starting time of TMS measurement), a Chi-square test for categorical data (attention score) and LMMs for data related to individual MEPs (angle and distance error). Furthermore, we used LMMs to estimate PAS-induced changes of MEP size, their correlations with baseline MEP size, and intraclass correlations (ICCs). For these LMMs, the dependent variable was MEP size, which was log2-transformed to better fit the assumption of normally distributed residuals. In addition, these LMMs were adjusted for log2-transformed angle and distance error.

We built Model 1 to estimate PAS-induced effects on MEP size at each time point *after induction* (Post 1, Post 2 and Post 3) within each session. This LMM included time point (categorical), session, and their interaction. The random effects included subject specific random effects for each time point in each session separately. An unstructured covariance matrix for the random effects was used (Model 1a) and was tested against the more restrictive compound symmetry structure (Model 1b), using a likelihood ratio (LR) test. Subsequently, the LR test was also used to assess the main effects of fixed effects.

Model 2 was built to estimate PAS-induced effects *during PAS induction*. This LMM included stimulus number (continuous), session and their interaction. Stimulus number was regarded as continuous time variable, as stimulations were equally spaced by 4 seconds in all PAS experiments. The model included subject specific random effects for stimulus number and session interaction and session. The eventual model was selected in three steps using likelihood ratio tests. First, we started out with a model using both natural cubic splines for stimulus number with three degrees of freedom and an unstructured covariance matrix (Model 2a). Second, to investigate the correlation structure, we tested Model 2a against a model with a compound symmetry structure (Model 2b). Last, to test whether the relation between MEP size and stimulus number was non-linear, Model 2a was tested against a model with a linear fit (Model 2c), using a LR test. After model selection, the LR test was also used to determine the main effects of fixed effects.

As a measure of within-subject consistency we calculated ICCs from LMMs that included session as an additional nesting level in the random effects. For the ICC of PAS-induced effects after induction, fixed effects and subject specific random effects of time point (categorical) were used (Model 3). To estimate the ICC of PAS-induced effects during PAS-induction over time, fixed effects as well as subject specific slopes for stimulus number (continuous time variable) were included (Model 4). Since the models used to calculate ICCs contained random effects for the respective time variables, the variation partition method was used [24]. 95% confidence intervals (95%CIs) for each ICC were estimated using 500 bootstrap samples.

Results

Session characteristics

Thirty-eight individuals (22 women; median age 23, range 19-38; mean IQ 107±10SD) underwent two PAS sessions, which were spaced at least 1 week apart (median days between sessions was 14, IQR: 4). As displayed in Table 1, median starting time was significantly earlier in session 1 than in session

239 2, whereas both sessions did not differ in terms of baseline RMT, SI1mV or the level of attention
240 during PAS induction. Additionally, the angle error and distance error relative to the hotspot did not
241 differ between sessions either. The estimated means of both errors over each time point were minimal,
242 with the upper limit of the 95%CIs not exceeding 5° for the angle error and 2mm for the distance
243 error. There was, however, a significant effect of time point (including PAS induction) on both angle
244 error (LR(8) = 22.41, p = 0.004) and distance error (LR(8) = 18.55, p = 0.018).
245

Table 1. Session characteristics and comparisons.

Characteristic	Session 1	Session 2	Statistic	P-value
RMT at baseline, mean (±SD), %MSO	48.5 (±10.3)	48.3 (±9.1)	t(37) = 0.17	0.87
SI1mV, %MSO	61.4 (±14.4)	60.1 (±14.7)	t(37) = 1.10	0.28
Start time, median (IQR), hh:mm	12:44 (12:28-13:08)	15:23 (15:00-15:41)	Z=719	<0.0001
Attention score, n (%)			LR(2) = 3.74	0.15
<4 stimuli reported	10 (29)	14 (38)		
4 stimuli reported	10 (29)	15 (41)		
>4 stimuli reported	15 (43)	8 (22)		
Angle error, estimated mean [95%CI]*, degrees			LR(5) = 5.26**	0.39
Baseline	2.69 [1.75, 4.16]	1.67 [1.11, 2.51]		
Induction	2.95 [1.99, 4.35]	2.11 [1.47, 3.01]		
Post1	2.07 [1.39, 3.08]	1.86 [1.12, 3.08]		
Post2	2.23 [1.41, 3.54]	1.29 [0.89, 1.87]		
Post3	1.83 [1.15, 2.91]	1.09 [0.81, 1.46]		
Distance error, estimated mean [95%CI]*, mm			LR(5) = 3.85**	0.57
Baseline	1.01 [0.84, 1.23]	0.87 [0.68, 1.12]		
Induction	1.11 [0.95, 1.29]	0.91 [0.79, 1.04]		
Post1	1.10 [0.97, 1.26]	0.91 [0.79, 1.06]		
Post2	1.20 [0.96, 1.50]	1.30 [1.02, 1.65]		

Post3	1.31 [1.02, 1.68]	1.34 [1.03, 1.74]

* Estimated using a LMM with the log2 transformed error as dependent variable and time point, session and their interaction as fixed effects.
** Main effect of session estimated by comparing LMMs using a likelihood ratio test

PAS-induced effects post induction

We determined the PAS-induced effect on MEP size at each post-induction measurement in each session. After filtering out MEPs with a noisy baseline, 5212 MEPs out of the originally 6080 MEPs recorded could be used for statistical analysis, resulting in the exclusion of one complete session of one subject. Plotting the individual trajectories of MEP size after induction (Fig 2A) indicates that PAS-induced changes of MEP size were highly variable. It is, therefore, not surprising that a model with an unstructured covariance matrix provided a better fit than one with a compound symmetry matrix ($LR(34) = 214.73$, $p < 0.001$), and was therefore selected as the appropriate matrix to estimate PAS-induced effects in our data.

The estimated mean of baseline MEP size of session 1 (0.54 mV; 95%CI [0.43, 0.68]) did not differ from that of session 2 (0.61 mV; 95%CI [0.53, 0.71]) ($LR(4)=2.26$, $p=0.689$). MEP size changed significantly over time ($LR(6) = 16.23$; $p = 0.013$), which was mainly driven by a negative effect on MEP size in Post 3 in session 2 (Table 2), instead of a positive effect on MEP size as one would expect when performing PAS. PAS-induced effects did not differ between sessions, as the interaction between time point and session was not significant ($LR(3) = 1.93$; $p=0.586$). The absence of this interaction is reflected by the similar profile of MEP size over time illustrated in Fig 2B. There was a moderate positive correlation between individual baseline MEP size and the PAS-induced effect at each time point in session 1 ($r_{Post\ 1} = 0.68$, $r_{Post\ 2} = 0.57$, $r_{Post\ 3} = 0.52$) and a poor positive correlation in session 2 ($r_{Post\ 1} = 0.23$, $r_{Post\ 2} = 0.21$, $r_{Post\ 3} = 0.30$).

Table 2. Fixed effects of PAS induction on MEP size per post-induction time point and session estimated by linear mixed effect modelling.

	Session 1	Session 2

Variable	β, %	95%CI, %	t (5165)	P-value	β, %	95%CI, %	t (5165)	P-value
Post 1	+16.83	[-9.05, 50.07]	1.22	0.22	-3.99	[-26.58, 25.55]	-0.30	0.77
Post 2	-10.00	[-30.92, 17.24]	-0.78	0.43	-23.92	[-42.44, 0.55]	-1.92	0.05
Post 3	-14.84	[-35.65, 12.69]	-1.12	0.26	-31.65	[-47.92, -10.30]	-2.74	0.006

Fig 2. PAS-induced effects for session 1 and 2 separately.

Subjects underwent two identical PAS sessions spaced >1 week apart, with session 1 displayed in red and session 2 in blue. (A) The mean change in MEP size over time per session (black line) plotted over the individual line plots (colored lines). (B) The mean change in MEP size over time for both sessions, where dots represent means of individual medians and bars represent their standard error. (C) Linear regression lines through all MEPs during PAS induction per session (black lines) plotted over the linear regression lines through MEPs per individual (colored lines). (D) The change of MEP size over time during the PAS induction, with every dot representing the mean MEP size over all participants for that stimulation number. Lines are fitted linear regression lines per session.

Potentiation during PAS induction

We, additionally, took a novel approach to determine the PAS-induced effect on MEP size by assessing this effect during PAS induction next to post induction alone. For this analysis, 9360 MEPs were available out of the 15200 MEPs recorded, due to filtering out MEPs with a noisy baseline. The exclusion of these MEPs resulted in the exclusion of one entire session for 9 subjects and the exclusion of both sessions for 4 subjects. Viewing the individual trajectories of MEP size development again indicates that there was high inter-individual variability (Fig 2C), which is reflected by the superior fit of the model with an unstructured covariance matrix to one with a compound symmetry covariance matrix ($LR(8) = 525.31$, $p < 0.001$). The development of MEP size over time appeared to be linear (Fig 2D), supported by the fact that a model with a cubic fit was not superior to one with a linear fit ($LR(4) = 2.69$, $p = 0.612$). Therefore, a model with an unstructured covariance matrix and a linear fit was selected as the appropriate model to estimate PAS-induced effects.

The estimated mean of MEP size at the start of PAS induction in session 1 (0.43 mV, 95%CI [0.27, 0.59]) did not differ from that in session 2 (0.44 mV, 95%CI [0.29, 0.66]) (LR(2) = 0.967, p = 0.617). There was a main effect of time (LR(1) = 13.36, $p < 0.001$), as a result of a significant positive increase of MEP size over time in both session 1 (+132%, 95%CI [+51%, +258%]) and session 2 (+79%, 95%CI [+19%, +169%]). However, there was no evidence of this time effect being different between sessions (LR(1) = 0.87, p = 0.35), reflected by the similar slope of the MEP size development in Fig 2D. There was a moderate negative correlation between MEP size at the start of PAS induction and the change in MEP size over time for session 1 (r = -0.51) and a weak negative correlation for session 2 (r = -0.41).

Consistency of PAS-induced effects

The within subject consistency of PAS-induced effects between the two sessions was small: Post 1 had an ICC of 0.08 (95%CI [0.03, 0.16]); Post 2 had an ICC of 0.07 (95%CI [0.03, 0.14]); and Post 3 had an ICC of 0.03 (95%CI [0.01, 0.09]) (Fig 3). Furthermore, the PAS-induced effects during induction showed a similarly poor within-subject consistency (ICC = 0.15; 95%CI [0.05, 0.35]) (Fig 3), despite their significant potentiation at group level. The ICC of baseline MEP size before induction was poor (ICC = 0.02; 95%CI [<0.01, 0.04]), as well as at the start of PAS induction (ICC = 0.24; 95%CI = [0.04, 0.42]). Nonetheless, the SI1mV did have a good within-subject consistency (ICC = 0.88; 95%CI [0.83, 0.96]), as did the RMT at different time points ($ICC_{Baseline}$ = 0.85, 95%CI [0.77, 0.92]; $ICC_{Post\ 1}$ = 0.83, 95%CI [0.79, 0.90]; $ICC_{Post\ 2}$ = 0.85, 95%CI = [0.79, 0.92]; $ICC_{Post\ 3}$ = 0.85, 95%CI [0.78, 0.92]).

Fig 3. Intra-individual correlation of PAS-induced effects.

Scatterplots per phase during or after induction illustrating the correlation between individual PAS-induced effects (dashed slopes) of session 1 against session 2. The slopes plotted are derived from the models used to calculate the ICC of PAS-induced effects within subject and between session.

Discussion

To our knowledge, we performed the largest study reported to date aimed at quantifying the within-subject consistency of PAS-induced effects. Additionally, this is the first PAS-TMS study that quantifies the change in MEP size during PAS induction over time and the first using a linear mixed model analysis approach. We performed two identical PAS sessions in one group of healthy volunteers, resulting in pronounced potentiation over time during PAS induction, which was not consistent within subjects. PAS-effects after induction did not show the expected potentiation, and these effects were not consistent within subjects either. Additionally, we demonstrated that a linear mixed model with an unstructured covariance matrix provides the best model fit for our PAS data.

PAS-induced effects during and after induction

The absence of post-induction potentiation in our study, and even the negative trend towards the last post-induction measurement, is certainly not in line with most PAS studies (for review see [11]), for which a few factors in our study could be responsible. First, our baseline MEP size is lower than the baseline of around 1mV shown in most PAS studies. In addition, we found a poor to moderate positive correlation between baseline MEP size and the PAS induced effects at each post-induction time point, suggesting that individuals with a low baseline were prone to show lower PAS-induced effects. Combined, these two observations could indicate that our stimulation intensity during induction was relatively lower than most PAS studies, possibly causing our PAS-induced effects to be smaller. Nonetheless, this hypothetical interaction between stimulation intensity and potentiation effect has never been specifically studied and relatively lower stimulation intensities that are still supra-threshold should still be able to potentiate MEP size, as double activation of the cortex by TMS and MNS is conserved. Second, coil placement during experiments could have been inaccurate. However, as angle and distance error of the coil position relative to the hotspot were minimal and carefully corrected for in our model, this is unlikely the case in our study. Last, the known compromising effect of sleepiness

346 on MEP size could be a contributor [25]. As our subjects remained seated and were not allowed to
347 perform any type of physical activity or specific types of mental activity between post-induction time
348 points, it is plausible that subjects became increasingly sleepy during this phase of the experiment.
349 Unfortunately, we cannot support this speculation with actual measures of sleepiness during
350 experiments, as these were not assessed.
351
352 Nevertheless, we did find a significant increase of MEP size during PAS induction that shows striking
353 resemblance to the increase in excitatory post synaptic potentials seen in STDP experiments in rodents
354 [26]. From these animal studies, we know that the potentiation during plasticity induction correlates
355 with the potentiation after induction. However, whether our observed potentiation of MEP size during
356 PAS induction is indeed a proxy for NMDA-dependent LTP as well, should be confirmed by studies
357 using a sham-stimulation control and preferably also placebo-controlled studies with an NMDA-
358 receptor antagonist. It is noteworthy, however, that in our study MEP size at the start of PAS induction
359 showed a negative correlation with PAS-induced effects during PAS induction. As MNS during paired
360 stimulations has a known acute inhibitory effect on MEP size, also known as short-latency afferent
361 inhibition [27, 28], lower MEP size at the start of induction could indicate successful paired
362 stimulations and, therefore, be related to a more prominent PAS-induced potentiation.
363

364 The advantage of using linear mixed models for PAS data

365 Our results provide insight in the potential advantage of LMMs for analyzing PAS data over
366 conventional analysis methods. Most importantly, we show that using an unstructured covariance
367 matrix provides a better model fit than using a compound symmetry matrix. First, our finding at least
368 demonstrates that PAS data can greatly benefit from a model of which the covariance matrix can be
369 flexibly designed and suggests that the compound symmetry matrix is unsuited for PAS data. As the
370 compound symmetry assumes that the correlations of MEP size between each time point are equal,
371 this result is actually not surprising for data sampled in a dynamic biological system such as the
372 human brain. Since conventionally used RM-ANOVA is restricted to a compound symmetry matrix,

373 LMMs are therefore more appropriate for the analysis of PAS data and are likely to deliver more
374 reliable results. Consequently, future consistent implementation of LMMs for analyzing PAS data will
375 improve the reproducibility between PAS studies.
376

377 # Consistency of PAS-induced effects

378 We show that PAS-induced effects, both during and after PAS induction, have poor intra-individual
379 consistency. However, one could argue that the lack of significant post-induction potentiation
380 compromises the validity of the consistency levels in this study. Nonetheless, we were able to show
381 significant potentiation during induction, which showed similar low consistency. Additionally, a lack
382 of significant potentiation at group level does not warrant that at individual level, as can be
383 appreciated from the individual traces in Fig 2A, and the low consistency found in this study indicates
384 that the subset of individuals showing potentiation were largely different in each session. It is,
385 therefore, unlikely that the consistency would be much higher in case of significant potentiation at
386 group level, which is in line with previous results reported by Fratello et al. (2006): in their study, they
387 also performed two identical PAS sessions within one group of healthy volunteers (n=18) and found
388 equally poor intra-individual consistency, despite finding significant PAS-induced potentiation both
389 PAS sessions [29].
390

391 The low consistency of PAS does not seem to hold for methods that measure brain plasticity in
392 humans in general, as intermittent theta burst stimulation (iTBS) was found to produce reasonably
393 consistent measures of brain plasticity with an ICC of 0.53 [30]. Another paradigm, anodal
394 transcranial direct current stimulation (aTDCS), showed moderate to no intra-individual consistency in
395 two separate studies [31, 32]. The contrast between the within-subject consistency between iTBS and
396 PAS can be understood from animal studies, suggesting that plasticity induction protocols based on
397 low frequency stimulations, such as STDP, are much more vulnerable to metaplasticity than high
398 frequency paradigms such as theta burst stimulation [33]. This is in line with findings in human TMS
399 studies, also reporting a vulnerability of PAS to metaplasticity [34, 35]. Altogether, this implies that

PAS is probably much more dependent on highly dynamic time-varying factors than iTBS, resulting in a lower within-subject consistency of PAS.

In our view, the inconsistency of PAS-induced plasticity has major implications for its use. As our results have revealed that PAS-induced effects cannot be seen as a trait, dividing subjects into PAS responders and PAS non-responders or even selecting subjects based on previous PAS results seems generally unsuitable. Moreover, our results suggest that PAS has limited value in longitudinal studies, for example, in studies monitoring drug effects over longer periods of time or in patient follow-up studies characterizing disease progression.

Conclusion

Our results demonstrate that PAS-induced effects have a high intra-individual variability and that PAS-induced potentiation of MEP size seems to be most pronounced at the end of the induction phase opposed to after induction. Additionally, our results support the use of LMMs for analyzing PAS data. The main implication of our results is that, due to its low consistency, PAS is unsuitable to track the development of human brain plasticity over longer periods of time.

Acknowledgements

We thank Dr. D. Rizopoulos for his helpful statistical advice, Dr. H.J. Boele for helping with developing the signal analysis software, Dr. H.G. van Steenis for technical support, and T. van Essen and D. Mani for their help preparing the data for signal analysis.

References

1. Caroni P, Donato F, Muller D. Structural plasticity upon learning: regulation and functions. Nat Rev Neurosci. 2012;13(7):478-90.

2. Caroni P, Chowdhury A, Lahr M. Synapse rearrangements upon learning: from divergent-sparse connectivity to dedicated sub-circuits. Trends Neurosci. 2014;37(10):604-14.
3. Klyubin I, Ondrejcak T, Hayes J, Cullen WK, Mably AJ, Walsh DM, et al. Neurotransmitter receptor and time dependence of the synaptic plasticity disrupting actions of Alzheimer's disease Abeta in vivo. Philos Trans R Soc Lond B Biol Sci. 2014;369(1633):20130147.
4. Srivastava AK, Schwartz CE. Intellectual disability and autism spectrum disorders: causal genes and molecular mechanisms. Neurosci Biobehav Rev. 2014;46 Pt 2:161-74.
5. Stefan K, Kunesch E, Cohen LG, Benecke R, Classen J. Induction of plasticity in the human motor cortex by paired associative stimulation. Brain. 2000;123 Pt 3:572-84.
6. Wolters A, Sandbrink F, Schlottmann A, Kunesch E, Stefan K, Cohen LG, et al. A temporally asymmetric Hebbian rule governing plasticity in the human motor cortex. J Neurophysiol. 2003;89(5):2339-45.
7. Stefan K, Kunesch E, Benecke R, Cohen LG, Classen J. Mechanisms of enhancement of human motor cortex excitability induced by interventional paired associative stimulation. J Physiol. 2002;543(Pt 2):699-708.
8. Wischnewski M, Schutter DJ. Efficacy and time course of paired associative stimulation in cortical plasticity: Implications for neuropsychiatry. Clin Neurophysiol. 2016;127(1):732-9.
9. Lahr J, Passmann S, List J, Vach W, Floel A, Kloppel S. Effects of Different Analysis Strategies on Paired Associative Stimulation. A Pooled Data Analysis from Three Research Labs. PLoS One. 2016;11(5):e0154880.
10. Ridding MC, Ziemann U. Determinants of the induction of cortical plasticity by non-invasive brain stimulation in healthy subjects. The Journal of physiology. 2010;588(Pt 13):2291-304.
11. Wischnewski M, Schutter DJLG. Efficacy and time course of paired associative stimulation in cortical plasticity: Implications for neuropsychiatry. Clinical Neurophysiology. 2016;127(1):732-9.
12. Aarts E, Dolan CV, Verhage M, van der Sluis S. Multilevel analysis quantifies variation in the experimental effect while optimizing power and preventing false positives. BMC Neurosci. 2015;16(1):94-.
13. Aarts E, Verhage M, Veenvliet JV, Dolan CV, van der Sluis S. A solution to dependency: using multilevel analysis to accommodate nested data. Nature neuroscience. 2014;17(4):491-6.
14. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971;9(1):97-113.
15. Rossi S, Hallett M, Rossini PM, Pascual-Leone A, Safety of TMSCG. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol. 2009;120(12):2008-39.
16. Rossi S, Hallett M, Rossini PM, Pascual-Leone A. Screening questionnaire before TMS: an update. Clin Neurophysiol. 2011;122(8):1686.
17. Wechsler D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corporation; 1999.
18. Axelrod BN. Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. Assessment. 2002;9(1):17-23.
19. Sale MV, Ridding MC, Nordstrom MA. Factors influencing the magnitude and reproducibility of corticomotor excitability changes induced by paired associative stimulation. Exp Brain Res. 2007;181(4):615-26.
20. Awiszus F. TMS and threshold hunting. Suppl Clin Neurophysiol. 2003;56:13-23.

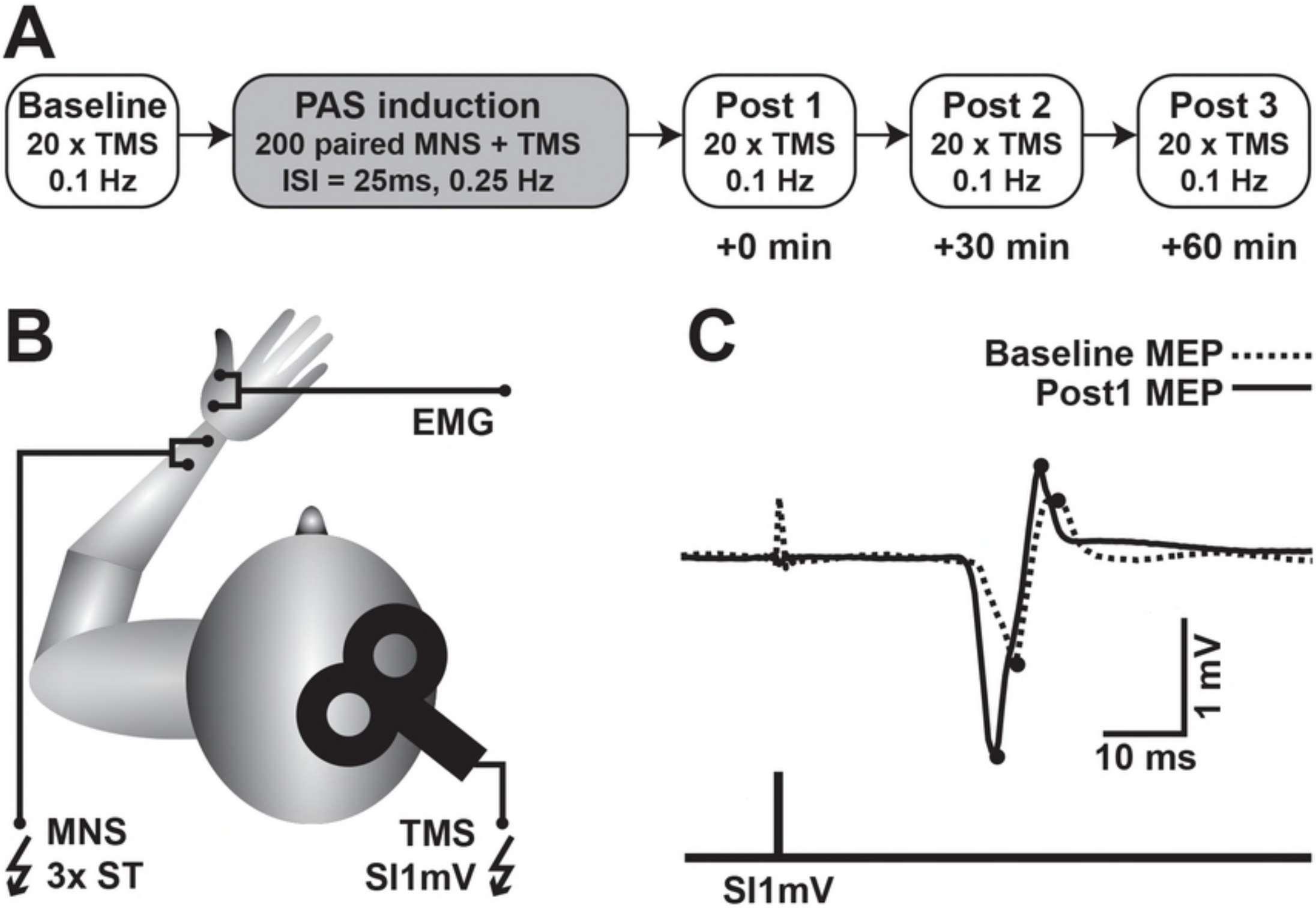
21. Stefan K, Wycislo M, Classen J. Modulation of associative human motor cortical plasticity by attention. J Neurophysiol. 2004;92(1):66-72.
22. R Development Core Team. Vienna, Austria: R Foundation for Statistical Computing; 2017.
23. Pinheiro J BD, DebRoy S, Sarkar D and R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. 2017.
24. Goldstein H, Browne W, Rasbash J. Partitioning Variation in Multilevel Models. Understanding Statistics. 2002;1(4):223-31.
25. Manganotti P, Fuggetta G, Fiaschi A. Changes of motor cortical excitability in human subjects from wakefulness to early stages of sleep: A combined transcranial magnetic stimulation and electroencephalographic study. Neuroscience Letters. 2004;362(1):31-4.
26. Froemke RC, Debanne D, Bi G-Q. Temporal modulation of spike-timing-dependent plasticity. Frontiers in synaptic neuroscience. 2010;2(June):19-.
27. Tokimura H, Di Lazzaro V, Tokimura Y, Oliviero A, Profice P, Insola A, et al. Short latency inhibition of human hand motor cortex by somatosensory input from the hand. J Physiol. 2000;523 Pt 2:503-13.
28. Turco CV, El-Sayes J, Savoie MJ, Fassett HJ, Locke MB, Nelson AJ. Short- and long-latency afferent inhibition; uses, mechanisms and influencing factors. Brain Stimul. 2018;11(1):59-74.
29. Fratello F, Veniero D, Curcio G, Ferrara M, Marzano C, Moroni F, et al. Modulation of corticospinal excitability by paired associative stimulation: reproducibility of effects and intraindividual reliability. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology. 2006;117(12):2667-74.
30. Hinder MR, Goss EL, Fujiyama H, Canty AJ, Garry MI, Rodger J, et al. Inter- and intra-individual variability following intermittent theta burst stimulation: Implications for rehabilitation and recovery. Brain Stimulation. 2014;7(3):365-71.
31. Lopez-Alonso V, Fernandez-Del-Olmo M, Costantini A, Gonzalez-Henriquez JJ, Cheeran B. Intra-individual variability in the response to anodal transcranial direct current stimulation. Clin Neurophysiol. 2015;126(12):2342-7.
32. Chew T, Ho K-A, Loo CK. Inter- and intra-individual variability in response to transcranial direct current stimulation (tDCS) at varying current intensities. Brain Stimulation. 2015;8(6):1130-7.
33. Zhang L. Hippocampal Synaptic Metaplasticity Requires Inhibitory Autophosphorylation of Ca2+/Calmodulin-Dependent Kinase II. Journal of Neuroscience. 2005;25(33):7697-707.
34. Delvendahl I, Jung NH, Mainberger F, Kuhnke NG, Cronjaeger M, Mall V. Occlusion of bidirectional plasticity by preceding low-frequency stimulation in the human motor cortex. Clin Neurophysiol. 2010;121(4):594-602.
35. Muller JF, Orekhov Y, Liu Y, Ziemann U. Homeostatic plasticity in human motor cortex demonstrated by two consecutive sessions of paired associative stimulation. Eur J Neurosci. 2007;25(11):3461-8.

Supporting Information

S1 Table. Reasons for excluding recruited individuals before entering the study.

520 **S1 Dataset. Full dataset of all variables reported.** RDA-file containing all data on all variables
521 reported in this manuscript. This file can be opened and used with the open source R or RStudio
522 software [22].

523 **S1 Syntax. Syntax for all linear mixed models reported.** Syntax written in the R-language,
524 containing all linear mixed models that are used to generate results for this manuscript. This file can
525 be opened and used with the open source R or RStudio software [22].


A
Baseline
20 x TMS
0.1 Hz
PAS induction
200 paired MNS + TMS
ISI = 25ms, 0.25 Hz
Post 1
20 x TMS
0.1 Hz
Post 2
20 x TMS
0.1 Hz
Post 3
20 x TMS
0.1 Hz
+0 min
+30 min
+60 min
B
EMG
MNS
3x ST
TMS
SI1mV
C
Baseline MEP
Post1 MEP
1 mV
10 ms
SI1mV

A

Session 1
Session 2

MEP size (mV)
3.0
2.0
1.0
.0

Baseline
Post 1
Post 2
Post 3
Baseline
Post 1
Post 2
Post 3

B

MEP size (mV)
1.0
.5
.0

Baseline
Post 1
Post 2
Post 3

C

Log2 MEP size (mV)
10
5
0

0
100
200
0
100
200

PAS induction stim. number

D

MEP size (mV)
2.0
1.0
.0

0
100
200

PAS induction stim. number

Post 1

Post 2

Post 3

Induction

Slope Session 2

Slope Session 1

4
3
2
1
0
-1
-2
-3

-3 -2 -1 0 1 2 3 4