Appendix

Supplemental Material

Augmented-feedback training improves cognitive motor performance of soccer players

Methods

Passing performance: coaches and players' judgments

Coaches were asked to judge the passing performance level of all tested players. Players and coaches were also asked to judge potential improvements of players. Importantly, they were both told that their judgments had to be based on a passing situation on the pitch, where a particular player would have to pass the ball to a moving teammate running 5 to 10 meters away from him (the closest situation to the passing test designed for the present study). The following procedures were adapted from our previous study (see 1 for details)

Passing performance judgments (coaches): Before providing them any information about the performance of players during the *passing* tests, coaches were asked to assess four aspects of the passing performance of every player. This was done through individual interviews between coaches and the same experimenter. A questionnaire had to be filled by each coach and the role of the experimenter was to explain the assessment procedure and instructions to coaches. For every line (player) of the questionnaire table, coaches had to use 3 graduated 5-point horizontal scales to assess, from *low* to *high*, the reactiveness (RE), the passing accuracy (PA) and the passing speed/power (PS, see 1 for details). After each judgment, coaches had to tell the degree of certainty in their judgments on another scale ranging from complete uncertainty (0 %) to complete certainty (100 %), using another graduated scale placed below each 5-point scale. Each coach filled this questionnaire few days before the beginning of the PRE and, for consistency checking, few days after the end of the POST training sessions, respectively. We thus collected a total of 324 judgments (6 coaches x 27 players x 2 PRE and POST judgments). Importantly, coaches were informed of the presence of three groups during the study but they were not informed to which group a player belonged to. However, they could easily infer which players belonged to the CON-group (who did not follow the additional visuomotor training protocol). In order to minimize the potential bias of this on the judgments, coaches were not allowed to check PRE scores when filling POST questionnaires (in other words, we did not provide them with a PRE reference for each player).

Passing performance - perceived evolution (coaches and players):

We used the same type of questionnaires to assess the perceived evolution of the passing performance by both players and coaches. Here, each player had to judge his own progress while coaches had to provide judgments for all players: this was done after the POST training sessions. Importantly, they had to consider the same situation of delivering a short-pass to a running teammate on the pitch. Rather than judging passing performance per se, they had to indicate the presence or absence of improvements in the passing performance using the same 5-point scale but now assessing from *not* to *much improved*. They were also asked to tell what potential cause could explain the perceived change by selecting one of four possibilities ('No idea', 'Field' training, 'Cognifoot' training or 'Both Cognifoot and Field' training). Differently from the previous questionnaire, coaches could here bias their judgments by taking into account the fact that CON-group players did not follow the visuomotor training.

Passing performance: objective measurements vs subjective judgments

Objective measurements of the passing performance were compared to coaches' judgments. For this purpose, COGNIFOOT measurements (RT, PSE and PS) were converted into RE_{score},

PA_{score} and PS_{score} using the same 5-point scales used by coaches (see 1 for details). The GPP_{score} was computed as (PA_{score} + RE_{score}) / 2. The evolution of the passing performance measured by COGNIFOOT (POST minus PRE scores) was compared to COACHES scores obtained using the two types of questionnaires mentioned in the previous section. For players, we also examined the perceived evolution of the performance by processing the scores obtained via the *perceived evolution* questionnaires.

Statistical analysis of coaches and players' judgments

In our previous study (1), we tested players of different ages (11-16 years of age) and observed that the correlation between COGNIFOOT and individual COACHES (PA/PS) scores was weak but became significant when COACHES scores were averaged. In particular, coaches could not judge accurately individual players' passing performances (a high dispersion of the coaches' data for a particular age was observed) but their judgments were followed an age-related linear effect on the passing performance, as detected by COGNIFOOT measurements.

Here, we only tested players of one age group but also observed weak correlations when comparing COGNIFOOT to COACHES scores. We therefore focused on the ability of coaches to detect a potential global effect of training on passing performance, and whether this effect corresponded to training related-performance changes physically measured by COGNIFOOT. First, we tested the internal consistency of coaches' judgments. This was measured using the ω_h coefficient (2, 3): a value of ω_h equal to or above 0.7 indicates that scores are coherent across coaches (which would then validate the computation of mean COACHES scores). We then performed ANOVAs to compare the effects of the training group on the perceived performance changes across PRE and POST sessions.

Results

Effects of training on passing performance

The passing performance was measured for different visuo-motor conditions: different visual (a visual target was moving in two directions, with two different speeds; the presence or absence of distractors – zero, one or two visual distractors moving agonistically/antagonistically relatively to the target motion) and different types of motor skills (eccentric passes or passes oriented towards the central part of the screen). In addition to the main effects of training on passing performance parameters (described in the main manuscript), all statistically significant interaction effects (training group x testing session x visuo-motor conditions) are detailed here.

Response times*:* RT were significantly different across categories of passes ($F(3, 48)=10.4$, $p<0.01$, $\eta_p^2=0.30$ - figure 1sm-A). Planned contrasts (*eccentric passes* vs *passes towards the center*) revealed that RT were shorter for eccentric passes (853 ± 92 and 856 ± 91 ms for PR-VR and PL-VL passes *vs* 884 ± 95 and 892 ± 90 ms for PC-VL and PC-VR passes, respectively, $t(48)=-5.14$, $p<0.01$). RT were also significantly shorter ($F(1, 24)=11.2$, $p<0.01$, $\eta_p^2=0.32$) for fast target speeds (881 ± 90 and 862 ± 87 for *moderate* and *fast* speeds, respectively). We observed a *PRE/POST training* x *pass category interaction effect* ($F(3, 72)=3.11$, $p=0.03$, $\eta_p^2=0.11$ - figure 1sm-A). Planned contrasts (*PRE* vs *POST* and eccentric passes vs passes towards the center) revealed that the PRE/POST difference in RT was significantly larger, within the category of passes towards the center, for PC-VL compared to PC-VR ($t(72)=-2.08$, $p=0.048$).

Passing spatial error: We observed a significant effect of the pass category ($F(3, 48)=3.44$, $p=0.024$, $\eta_p^2=0.18$). Planned contrasts (eccentric passes vs passes towards the center) revealed that passes were more accurate for targets towards the center compared to eccentric passes (46.1 ± 31.4 and 41.0 ± 29.7 cm for PR-VR and PL-VL passes *vs* 38.5 ± 30.5 and 37.6 ± 33.7 ms for PC-VL and PC-VR passes, respectively, $t(48)=8.71$, $p<0.01$). We also observed a tendency for larger PSE for the *fast* speed (37.8 ± 6.6 and 43.9 ± 10.0 cm for *moderate* and *fast* speeds, respectively). However, this effect was not confirmed statistically although it was close to significance ($F(1, 16)=4.08$, $p=0.06$, $\eta_p^2=0.20$). The following interaction effects were found to be statistically significant (because Levene's test of homogeneity of variances revealed that variances were unequal only when CON-group was included - $F(2, 24)=3.93$, $p=0.03$, the following comparisons were performed only for AF-group and NF-group, see main manuscript):

- *Pass Category* x *Target Speed* ($F(3, 48)=2.87$, $p=0.046$, $\eta_p^2=0.15$, figure 2sm-A). Planned contrasts (*fast* vs *moderate* and eccentric passes vs passes towards the center) revealed that the PSE difference across speeds is significantly larger for eccentric passes ($t(48)=6.36$, $p=0.022$).

- *Pass Category* x *Target Speed* x *PRE/POST training* ($F(3, 48)=4.05$, $p=0.012$, $\eta_p^2=0.20$ - figure 2sm-B). Planned contrasts (PRE *vs* POST / eccentric passes *vs* passes towards the center / moderate *vs* fast target speeds), revealed a difference in the evolution of PSE across speeds depending on the category of passes ($t(48)=10.26$, $p<0.01$). We therefore tested specific contrasts for moderate or fast speeds. For moderate speeds, planned contrasts on categories of passes (PC-VL vs the three other categories) revealed that PSE decreased significantly more after training for PC-VL passes ($t(48)=4.97$, $p=0.04$, figure 2sm-B, left panel). For fast speeds, planned contrasts (eccentric passes vs passes towards the center) revealed that PSE decreased with the same magnitude after training irrespective of the category of passes (p>0.05) although this decrease visually seemed to be more important for eccentric passes (figure 2sm-B, right panel).

- *PRE/POST training* x *Pass Category* x *Group* (F(3, 48)=3.50, p=0.02, η_p^2 =0.18, figure 2sm-C). Planned contrasts (PRE *vs* POST / eccentric passes *vs* passes towards the center / AF-group *vs* NF-group), revealed that PSE is significantly smaller after training only for passes towards the center in the NF-group (t(48)=4.79, p=0.043) while no statistically significant effect of the pass category was observed in the AF-group (p>0.05).

Passing speed: On average, PS increased by 1.7 km/h and decreased by 1.6 and 1.4 km/h after training, for the AF-group, NF-group and CON-group, respectively. We observed only a statistically significant effect of the *target speed* on PS (39.7 ± 3.7 and 41.1 ± 3.9 for *moderate* and *fast* speeds, respectively; F(1, 24)=15.3, p<0.01, η_p^2 =0.39).

Global Passing Performance: A statistically significant *PRE/POST x Pass Category* interaction effect (F(3, 72)=3.76, p=0.014, η_p^2 =0.13 - figure 1sm-B) followed by planned contrasts (*PRE* vs *POST* and PC-VL passes vs the three other types of passes) revealed that the *PRE*/*POST* difference in GPP was larger for PC-VL passes (t(72)=24.2; p<0.001), the PRE/POST difference in GPP being comparable when comparing the other categories of passes (p>05). A statistically significant *PRE/POST x Pass Category* interaction effect (F(3, 72)=3.76, p=0.014, η_p^2 =0.13 - figure 1sm-B) followed by planned contrasts (*PRE* vs *POST* and PC-VL passes vs the three other types of passes) revealed that the *PRE*/*POST* difference in GPP was larger for PC-VL passes ($t(72)=24.2$; $p<0.001$), the PRE/POST difference in GPP being comparable when comparing the other categories of passes ($p>05$).

Real versus *perceived* evolution of the passing performance in coaches

Measure of a potential bias on coaches' judgments

As mentioned in the Methods section of the main manuscript, coaches were informed of the presence of three groups during the study but they were not informed to which particular group a player belonged to. However, they could easily infer which players belonged to the CON-group (who did not follow the training protocol). This bias would make coaches providing lower POST scores in the CON- group in particular (as observed in figure 4 of the main manuscript). However, we can exclude this possibility for two reasons.

First, in order to minimize this potential bias, coaches were not allowed to check the PRE scores when filling the POST questionnaires (in other words, we did not provide them with a PRE reference for each player).

Second, coaches filled both POST and 'perceived evolution' questionnaires the same day (they first filled the POST questionnaire). Any bias in the direction of lower 'performance improvement' scores for the CON-group should be observed in the 'perceived evolution' questionnaires. We compared these scores across groups and coaches using ANOVAs and did not observe any effect of the *training group* on the evaluation of the Global Passing Performance improvement ($p>0.05$). Similarly, no *training group* x *coaches* interaction effect was observed ($p>0.05$). Furthermore, the Bayes factor in favor of the null hypothesis was equal to 5.65, indicating that the absence of any effect of the *training group* on the evaluation of the passing performance improvement.

This demonstrates that coaches provided their judgments on the performance of players, independently of the *training group* to which players belonged to.

Coherence of the coaches' judgments

As detailed in the Methods section, the coherence of coaches' judgments was measured using the ω_h coefficient (2, 3). The judgments provided by coaches during the PRE ($\omega_{h.PA}$ = 0.83, $\omega_{h.PS}$ = 0.86, $\omega_{h.RE}$ = 0.77 and $\omega_{h.GPP}$ = 0.86) and POST sessions ($\omega_{h.PA}$ = 0.87, $\omega_{h.PS}$ = 0.85, $\omega_{h.RE}$ = 0.86 and $\omega_{h.GPP}$ = 0.90) were well above the $\omega_h > 0.7$ coherence criterion. In contrast, the scores provided by coaches in the *perceived evolution* questionnaires were not coherent ($\omega_{h.PA}$ = 0.19, $\omega_{h.PS}$ = 0.53, $\omega_{h.RE}$ = 0.33 and $\omega_{h.GPP}$ = 0.30). Based on these observations, we computed the mean evolution of performance based on PRE and POST questionnaires.

Perceived cause of the performance evolution and degree of certainty (coaches)

The perceived cause for performance change is a variable following an ordinal scale ('No idea', 'Field' training, 'Cognifoot' training or 'Both Cognifoot and Field' training). These data are presented in figure 3sm-A2/B2/C2. We run χ^2 tests to examine whether the distribution of causes were affected by the *training group*. We also tested how the training group affected the degree of certainty of coaches' judgments (data presented in figure 3sm-A3/B3/C3). Note here that since GPP scores were computed from PA and RE scores (see Methods), the analysis of the causes of GPP improvement could not be performed.

<u>Reactiveness:</u> Coaches judged that 'Cognifoot training' contributed less to RE scores changes following training in CON-group compared to AF-group /NF-group (figure 3sm-A2): around 65 % and 20 % of AF-group /NF-group and CON-group players were judged as having improved RE because of 'Cognifoot' or 'Cognifoot + Field' training, respectively. Chi-square tests revealed a significant effect of the *training group* ($\chi^2(6) = 56.4$, $p<0.001$), with a

significant effect between AF-group and CON-group ($\chi^2(3) = 43.3$, $p<0.001$) and between NF-group and CON-group ($\chi^2(3) = 40.2$, $p<0.001$), no difference being observed between AF-group and NF-group ($p>0.05$). The degree of certainty in RE judgments did not significantly differ across training groups ($p>0.05$, figure 3sm-A3).

Passing accuracy: The distribution of causes of PA score changes after training (figure 3sm-B2) was similar to RE (figure 3sm-B2). This was confirmed by a significant effect of *training group* on the distribution of causes ($\chi^2(6) = 60.7$, $p<0.001$), with a significant effect between AF-group and CON-group ($\chi^2(3) = 41.0$, $p<0.001$) and between NF-group and CON-group ($\chi^2(3) = 46.3$, $p<0.001$), no difference being observed between AF-group and NF-group ($p>0.05$). ANOVA revealed that the degree of certainty in PA judgments did not differ across training groups ($p>0.05$, figure 3sm-B3).

Passing speed: The distribution of causes of PS score changes following training (figure 3sm-C2) was similar to PA and RE. Chi-square tests revealed a significant effect of the *training group* ($\chi^2(6) = 70.4$, $p<0.001$), with a significant effect between AF-group and CON-group ($\chi^2(3) = 51.3$, $p<0.001$) and between NF-group and CON-group ($\chi^2(3) = 25.0$, $p<0.001$), no difference being observed between AF-group and NF-group ($p>0.05$). The degree of certainty in RE judgments did not significantly differ across training groups ($p>0.05$, figure 3sm-C3).

Taken together, these results show that coaches judged that Cognifoot-training was involved in the performance improvements in 65 % of AF-group and NF-group players.

Real versus *perceived* evolution of the passing performance in players

Perceived evolution of the performance (players)

Reactiveness: No significant effect of the *training group* ($p>0.05$) was observed on players' $RE_{evolution}$ scores (figure 3sm-A1).

Passing accuracy: No significant effect of the *training group* ($p>0.05$) was observed on players' $PA_{evolution}$ scores (figure 3sm-B1). However, planned contrasts (AF-group /CON-group vs NF-group) revealed that the perceived PA score improvement was significantly lower ($t(24) = 2.10$, $p=0.046$) in NF-group (around 50 %) compared to AF-group and CON-group players (around 62 and 70 %, respectively), no difference being observed between AF-group and CON-group ($p>0.05$).

Passing speed: No significant effect of the *training group* ($p>0.05$) was observed on players' $PS_{evolution}$ scores (figure 3sm-C1) although a tendency for higher values can be observed in AF-group players.

Global passing performance: No significant effect of the *training group* ($p>0.05$) was observed on players' $GPP_{evolution}$ scores (figure 3sm-D) although a tendency for lower values can be observed in NF-group players.

Taken together, these data imply that players hugely over-estimated their performance improvements compared to COGNIFOOT and coaches' judgments (see figure 3sm).

Perceived cause of the performance evolution and degree of certainty (players)

Reactiveness: Compared to CON-group, AF-group and NF-group players judged that 'Cognifoot training' contributed more to RE scores changes (figure 3sm-A2): 100 % / 80 % and 10 % of AF-group /NF-group and CON-group players judged that 'Cognifoot' was the main cause for the RE score improvement they reported, respectively. Interestingly, while they did not follow any COGNIFOOT-based training, around 90 % of CON-group players judged that ‘Cognifoot + Field’ training caused them to improve RE. When asked them to explain this contradiction, CON-group players wrote that the first PRE test in which they participated helped them to focus more on the quality of their passes to running teammates in the following field training sessions. It is noticeable that neither physical measurements nor coaches’ judgments indicated any RE improvement in NF-group players (figure 3sm-A). Chi-square tests revealed a significant effect of the *training group* on the distribution of causes on performance changes ($\chi^2(4) = 20.8$, $p < 0.001$; note that the number of degrees of freedom is equal to 4 because not all causes were present in the different groups). A significant effect between AF-group and CON-group ($\chi^2(1) = 14.4$, $p < 0.001$) and between NF-group and CON-group was observed ($\chi^2(2) = 10.9$, $p < 0.01$), while no difference was detected between the AF-group and NF-group ($p>0.05$).

ANOVA revealed a significant effect ($F(2, 24)=5.24$, $p=0.013$) of the training group on the degree of certainty in $RE_{evolution}$ judgments (figure 3sm-A3). Planned contrasts (AF-group /CON-group vs NF-group) revealed that the degree of certainty was significantly lower in NF-group than in AF-group /CON-group groups (by up to 20 %; $t(24) = 3.22$, $p < 0.01$), no difference being observed AF-group and CON-group ($p > 0.05$).

Passing accuracy: Players of the AF-group /NF-group judged that ‘Cognifoot training’ contributed more to PA scores changes than CON-group players (figure 3sm-B2): around 65 % / 30 % and 0 % of AF-group /NF-group and CON-group players judged that ‘Cognifoot’ was the main cause for the PA score improvement they reported, respectively. Here also, 55 % of CON-group players judged that ‘Cognifoot + Field’ training caused them to improve PA although neither physical measurements nor coaches indicated PA improvement in CON-group players (figure 3sm-B). Chi-square tests did not reveal any significant effect of the *training group* ($p > 0.05$) on the distribution of causes of performance changes. A significant effect between AF-group and CON-group was observed ($\chi^2(2) = 9.3$, $p = 0.026$) but there were no differences between the other groups ($p>0.05$).

ANOVA revealed that the degree of certainty in $PA_{evolution}$ judgments (figure 3sm-B3) did not significantly differ across training groups although the p value was close to the significance level ($F(2, 24)=2.86$, $p=0.076$). Planned contrasts (AF-group /CON-group vs NF-group) revealed that the degree of certainty of NF-group players (around 63 %) was significantly ($t(24) = 2.38$, $p = 0.025$) lower than the one of AF-group /CON-group (around 80 %), no difference being observed between AF-group and CON-group ($p > 0.05$).

Passing speed: Players of all groups judged that 'Cognifoot' and 'Cognifoot + Field' training contributed to $PS_{evolution}$ scores improvements (from 66 % for the CON-group to 100 % in the AF-group, figure 3sm-C2). Chi-square tests did not reveal any significant effect of the *training group* on the distribution of causes on performance changes ($p>0.05$). Interestingly, this perception of improved PS score was observed in coaches' scores but not physically measured by COGNIFOOT (figure 3sm-C).

ANOVA revealed a significant effect ($F(2, 24)=5.0$, $p=0.015$) of the training group on the degree of certainty related to $RE_{evolution}$ judgments (figure 3sm-C3). Planned contrasts (AF-group /CON-group vs NF-group) revealed that the degree of certainty of NF-group players was significantly lower (by up to 20 %; $t(24) = 2.96$, $p < 0.01$) than the one of AF-group /CON-group, no difference being observed between AF-group and CON-group ($p > 0.05$).

Summary

Overall, we noticed that both coaches and players perceived significant changes of passing performance following training. However, players' scores seem to be largely over-estimated (including players of the CON-group) while the physically-measured effect of the *training group* on the passing performance was noticed in the coaches' scores only (figure 3sm-A1/B1/D). Interestingly, coaches reported that Cognifoot-training was involved ('Cognifoot' or 'Cognifoot + Field' trainings) in this perceived improvement in 65 % of players of AF-group and NF-group (figure 3sm-A2/B2/C2). The contribution of Cognifoot – training to performance improvements judged by coaches was around 15-20 % in the CON-group (figure 3sm-A2/B2/C2, while CON-group players judged this contribution to be ranged between 60 % (PS score, figure 3sm-C2) and 90 % (RE score, figure 3sm-A2). Since no change in passing performance was noticed in CON-group players, we can conclude that coaches' judgments are more reliable than players' judgments.

1. Hicheur H, Chauvin A, Chassot S, Cheneviere X, Taube W. Effects of age on the soccer-specific cognitive-motor performance of elite young soccer players: Comparison between objective measurements and coaches' evaluation. *PloS one*. 2017;12(9):e0185460.

2. Revelle WZ, R.E. Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*. 2009;74(1):145.

3. Trizano-Hermosilla I, Alvarado JM. Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in psychology*. 2016;7:769.

Figure legends

Figure 1sm: **A-** Response times as a function of *Pass Category* and *PRE/POST training*, and **B-** Global Passing Performance as a function of *Pass Category* and *PRE/POST training*.

Figure 2sm: Passing spatial error as a function of **A-** *Pass Category* and *Target Speed* (PL and PR correspond to Passes towards the Left and Right –eccentric passes, respectively; PC correspond to Passes towards the center. VL and VR correspond to leftward and rightward target motion, respectively), **B-** *Pass Category*, *Target Speed* and *PRE/POST training* and **C-** *PRE/POST training*, *Pass Category* and *Group* (AF-group and NF-group correspond to the Augmented-Feedback and No-Feedback groups, respectively).

Figure 3sm: A1 to D1 - Evolution of the passing performance scores measured by COGNIFOOT or judged by COACHES / PLAYERS (RE- Reactiveness, PA-Passing accuracy, PS- Passing speed, and GPP- Global Passing Performance); the dashed grey horizontal line at y = 0 indicate the absence of performance improvement. A2 to C2 - Perceived causes of the changes in (RE - PA - PS) performance following training (COACHES and PLAYERS). A3 to C3 - Degrees of certainty in COACHES and PLAYERS' judgments (RE - PA - PS) across groups.

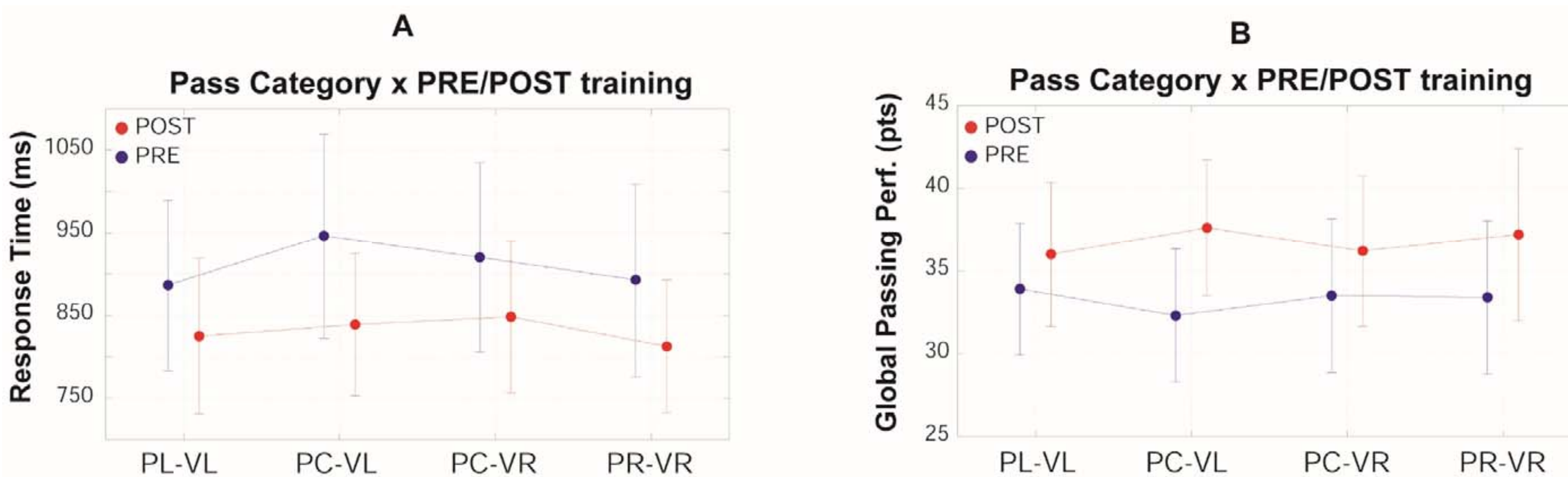
A
Pass Category x PRE/POST training
Response Time (ms)
1050
950
850
750
POST
PRE
PL-VL
PC-VL
PC-VR
PR-VR
B
Pass Category x PRE/POST training
Global Passing Perf. (pts)
45
40
35
30
25
POST
PRE
PL-VL
PC-VL
PC-VR
PR-VR

Figure 1sm

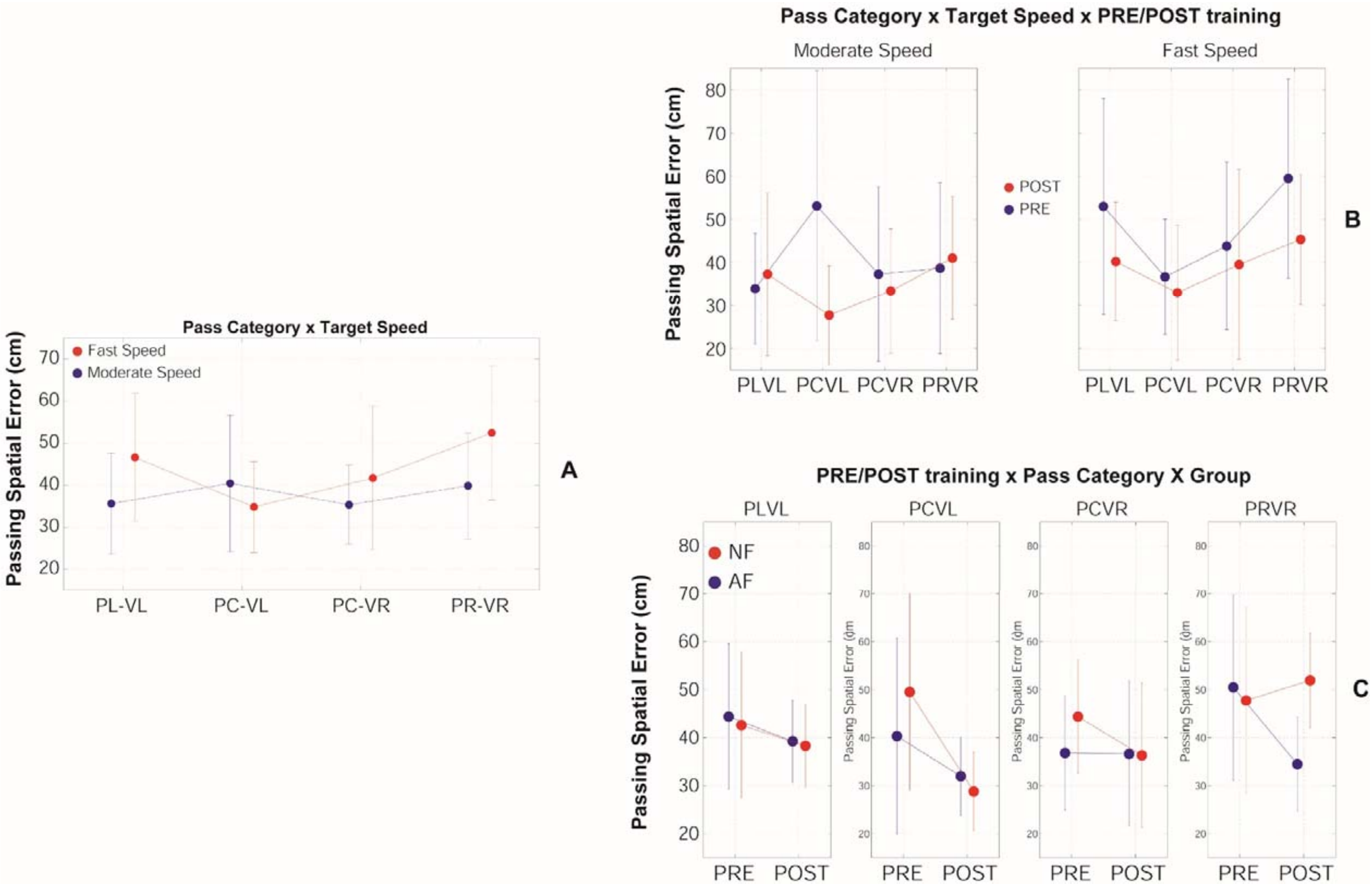
Pass Category x Target Speed
Fast Speed
Moderate Speed
Passing Spatial Error (cm)
PL-VL
PC-VL
PC-VR
PR-VR
A
Pass Category x Target Speed x PRE/POST training
Moderate Speed
Fast Speed
POST
PRE
PLVL
PCVL
PCVR
PRVR
B
PRE/POST training x Pass Category X Group
NF
AF
PRE
POST
C

Figure 2sm

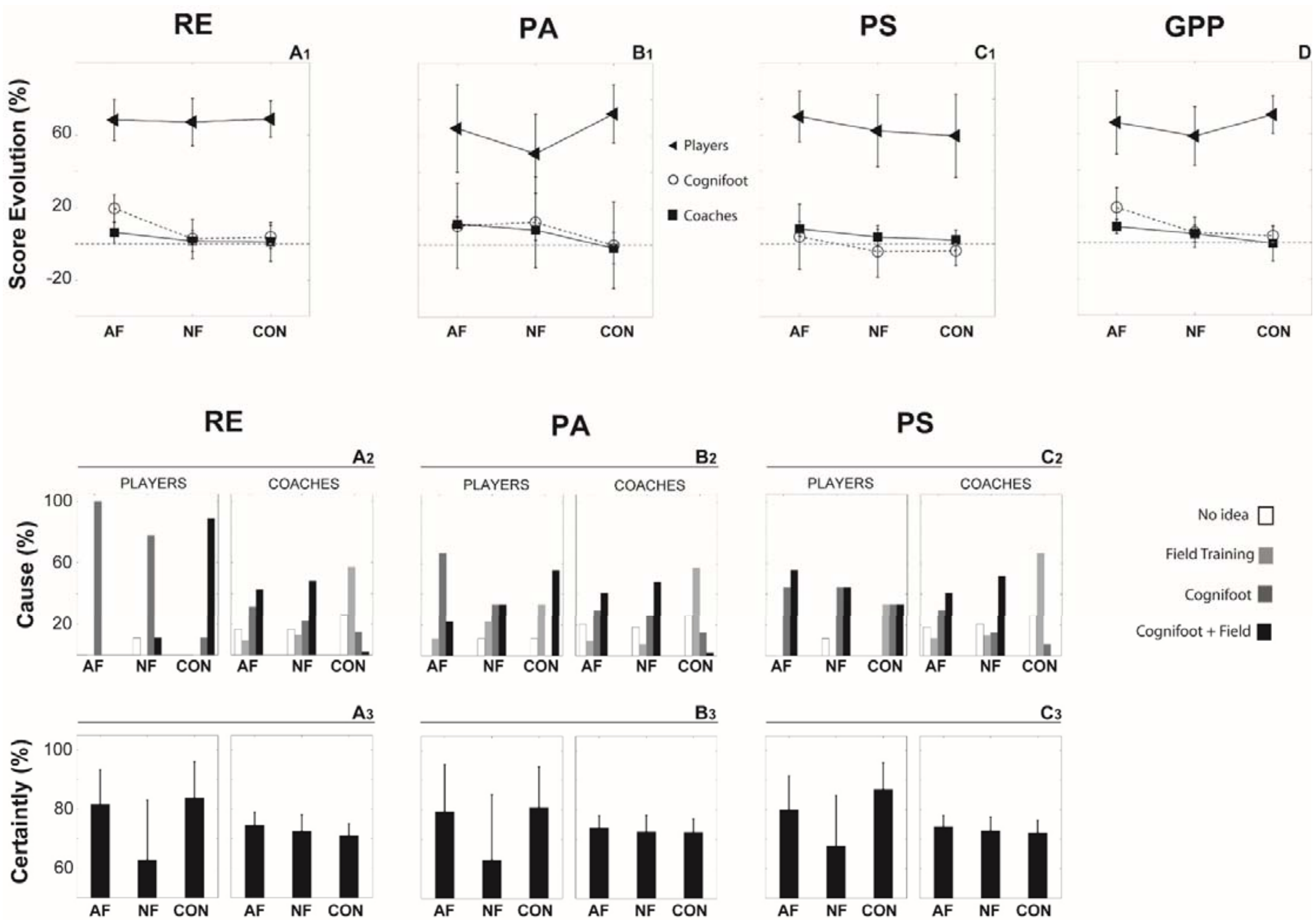
RE
PA
PS
GPP
A1
B1
C1
D
Score Evolution (%)
60
20
-20
AF
NF
CON
Players
Cognifoot
Coaches
A2
B2
C2
PLAYERS
COACHES
Cause (%)
100
No idea
Field Training
Cognifoot
Cognifoot + Field
A3
B3
C3
Certaintly (%)
80

Figure 3sm