Educational Research Review 18 (2016) 46–57

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis☆

Giovanni Sala*, Fernand Gobet

Department of Psychological Sciences, University of Liverpool, United Kingdom

ARTICLE INFO

Article history:
Received 21 October 2015
Received in revised form 20 January 2016
Accepted 25 February 2016
Available online 27 February 2016

Keywords:
Chess
Transfer
Meta-analysis
Learning
Education

ABSTRACT

In recent years, pupils' poor achievement in mathematics has been a concern in many Western countries. Chess instruction has been proposed as one way to remedy this state of affairs, as well as improving other academic topics such as reading and general cognitive abilities such as intelligence. The aim of this paper is to quantitatively evaluate the available empirical evidence that skills acquired during chess instruction in schools positively transfer to mathematics, reading and general cognitive skills. The selection criteria were satisfied by 24 studies (40 effect sizes), with 2788 young people in the chess condition and 2433 in the control groups. The results show (a) a moderate overall effect size ($g = 0.338$); (b) a tendency for a stronger effect on mathematical ($g = 0.382$) than reading skill ($g = 0.248$), and (c) a significant and positive effect of duration of treatment ($Q(1) = 3.89$, $b = 0.0038$, $p < .05$). However, no study used an “ideal design” including pre- and post-test, full random allocation of participants to conditions and, most importantly, both a do-nothing control group and an active control group - a problem common in education research. Directions for further research are discussed.

Contents

☆ We thank William Bart, Guillermo Campitelli, and Grace McGuire for useful comments on an earlier draft of this paper.

* Corresponding author. Department of Psychological Sciences, Bedford Street South, University of Liverpool, Liverpool, L69 7ZA, United Kingdom.
E-mail address: Giovanni.sala@liv.ac.uk (G. Sala).

http://dx.doi.org/10.1016/j.edurev.2016.02.002
1747-938X/

1. Introduction

Recently, many concerns have been expressed about pupils' poor mathematics achievement both in the United States (Hanushek, Peterson, & Woessmann, 2012; Richland, Stigler, & Holyoak, 2012) and in Europe (Grek, 2009). Pupils' low mathematical skills have serious consequences well beyond the classroom, as the possibility of successfully majoring in Science, Technology, Engineering, and Mathematics (STEM) subjects, and hence obtaining STEM jobs, is limited by one's mathematical skills. The job market demands more graduates in STEM subjects than graduates in the humanities and has also become more competitive worldwide in recent years, with increasingly high mathematical competences being required (Halpern et al., 2007).

To address the issue of how to improve mathematics instruction, policy makers and researchers have explored a number of avenues. One such avenue is to teach chess in schools. Chess has recently started to become part of the school curriculum (as an optional subject) in several countries. Chess-related research and educational projects are currently ongoing in the United Kingdom, Spain, Turkey, Germany, and Italy, among other countries. Commenting on a large project having introduced chess in the curriculum of 175 schools in the UK, chess master Jerry Myers stated that chess "directly contributes to academic performance. Chess makes children smarter" (Garner, 2012). The European Parliament has expressed its favourable opinion on using chess courses in schools as educational tool (Binev, Attard-Montalto, Deva, Mauro, & Takkula, 2011) and, similarly, the Spanish Parliament has approved the implementation of chess courses during school hours. These initiatives have been conducted because chess is considered an effective educational tool able to improve not only mathematical skills, but also other academic skills such as reading and general cognitive abilities such as concentration and intelligence, and even children's heuristics and habits of mind (Costa & Kallick, 2009). Critically, efforts to promote chess in schools take for granted that chess skill transfers to other domains.

1.1. Difficulty of transfer

Transfer of learning occurs when a set of skills acquired in one domain generalizes to other domains or improves general cognitive abilities. Transfer is an important question both theoretically and practically. Mestre (2005) distinguishes between *near-transfer*, where transfer occurs between closely related domains (e.g., transfer from geometry to calculus) and *far-transfer*, where the source and target domains are only loosely related (e.g., transfer from Latin to geography). It has been proposed that transfer is a function of the extent to which two domains share common features (Thorndike & Woodworth, 1901) and cognitive elements (Anderson, 1990). In line with this hypothesis, near-transfer is often observed, although exceptions also exist. For example, research into expertise shows that transfer is only partial between subspecialties of expertise such as cardiology and neurology (Rikers, Schmidt, & Boshuizen, 2002). By contrast, substantial research in education and psychology suggests that far-transfer is difficult (Donovan, Bransford, & Pellegrino, 1999). This includes the research on teaching the computer language LOGO in order to improve children's thinking skills, which has obtained disappointing results (De Corte & Verschaffel, 1986; Gurtner, Gex, Gobet, Nunez, & Restchitzki, 1990). In addition, the higher the level of a skill, the more specific the features of a domain will be, and the lower the likelihood that there will be transfer (Ericsson & Charness, 1994), in particular because a large number of domain-specific perceptual chunks will be acquired (Gobet, 2015). Again, there are exceptions, and some individuals have excelled in several different domains (Gobet, 2011; 2015).

The difficulty of transferring knowledge and skills raises a number of significant practical issues, especially in education. Most educational interventions try to transmit knowledge which, to some extent, is meant to be transferable from one domain of learning to another. In fact, transferability of skills is either a tacit assumption or a specific aim of nearly every educational program (Donovan et al., 1999; Perkins & Salomon, 1994). Therefore, educational institutions are interested in methodologies implementing school activities that teach and boost transferable skills. One approach is to teach general strategies, such as learning, problem-solving, and reasoning heuristics (Perkins & Grotzer, 2000), so that these skills can be easily transferred to other domains. Another approach is to teach a specific activity, with the hope that this activity will help individuals to develop skills that might be useable in other domains. The game of chess is one such activity that has been used in that way.

1.2. The issue of transfer in chess research

A substantial amount of research has been devoted to understanding the cognitive processes underpinning chess skill, and much is known about chess players' perception, learning, memory, and problem solving (for reviews, see Gobet, 2015; Gobet, De Voogt, & Restchitzki, 2004). Much less is known about the extent to which chess skill transfers to other domains of learning.

Several studies (Bilalić, McLeod, & Gobet, 2007; Doll & Mayr, 1987; Frydman & Lynn, 1992; Grabner, Stern, & Neubauer, 2007) have shown that chess players tend to be more intelligent than the general population. However, these studies were correlational in nature and cannot establish that chess skill is the actual cause of better intellectual abilities. In fact, the exact opposite causal explanation could be true: some individuals could excel at chess due to their superior intellectual abilities (Gobet & Campitelli, 2002).

Assuming that skills acquired in chess will lead to benefits in domains such as mathematics and reading clearly implies the presence of far transfer. In line with Thorndike and Woodworth's (1901) hypothesis, several studies have shown that chess players' skill tends to be context-bound, suggesting that it is difficult to achieve far-transfer from chess to other domains. For example, memory for chess positions fails to transfer from chess to digits both in adults and children (Chi, 1978; Schneider, Gruber, Gold, & Opwis, 1993); chess players' perceptual skills do not transfer to visual memory of shapes (Waters, Gobet, & Leyden, 2002); chess skill does not predict performance in the economic game known as beauty contest (Bühren & Frank, 2010); and finally, chess planning skills do not help chess players to solve the Tower of London task (Unterrainer, Kaller, Leonhart, & Rahm, 2011).

1.3. Chess in school

In spite of these negative results, several researchers have pursued the hypothesis that skills acquired with chess can transfer to other domains. Two main explanations have been adduced to support this hypothesis. First, chess requires decision-making skills and high-level processes (such as acquiring and selecting relevant information from a problem) similar to those used in mathematics and reading (Margulies, 1992). Second, since chess is a demanding task involving focused attention and problem solving, playing chess should strengthen these cognitive abilities and thus be beneficial for children's school performance (Bart, 2014). However, convincing experimental evidence of the effectiveness of chess instruction is lacking. In a literature review, Gobet and Campitelli (2006) argued that there was no strong evidence for the cognitive and academic benefits of chess. They found only few studies, which included unpublished reports or master and doctoral theses. Most importantly, many of these studies had a quasi-experimental design (no random assignment to the experimental and control groups) and, in some cases, the experimental samples were self-selected.

The difficulty of transferring chess skill is consistent with the literature on the transfer of specific skills. At first blush, it is hard to see why knowing the strategic value of the bishop pair or the correct way to handle a minority attack should offer any advantage in mathematics, understanding a text, or developing focused attention. Nevertheless, it is possible that chess practice enhances some abilities shared with other domains, such as those mentioned above, provided that chess is taught early on with children, when academic and cognitive abilities are at the beginning of their development. This is the reason why chess intervention studies have focused on the academic and cognitive skills of children rather than adults: Children's skills are less context-specific than adults', and thus transfer of learning is more likely in the former than in the latter.

Some recent studies (Sala, Gorini, & Pravettoni, 2015; Scholz et al., 2008; Trinchero, 2012; Trinchero & Sala, 2016) have provided more refined explanations as to why chess may effectively enhance cognitive and mathematical skills. According to these researchers, chess improves children's mathematical skills because the game has some elements in common with the mathematical domain and because it promotes suitable habits of mind (Costa & Kallick, 2009). Through chess, children train several context-independent skills (such as the ability to understand the existence of a problem or the need for correct reasoning), which may transfer to the mathematical domain. This is possible because (primary school) mathematics and chess share some common features (e.g., numerical and spatial relationships as well as quantity-based problems), strategies to solve problems (e.g., focusing and interpreting game/problem situations, selecting relevant information, or looking for correct arguments), cognitive skills (e.g., attention) and meta-cognitive skills (e.g., planning). The aim of our study is thus to test, comprehensively and quantitatively, these previous claims on the putative benefits of chess instruction in school.

2. Scope, aims, and hypotheses of the present meta-analysis

Given the considerable attention that research on chess in school is attracting and the potentially important implications for our understanding of transfer, it is important to provide a scientific evaluation of the effects of chess instruction on academic and cognitive skills. A similar interest has been devoted to studies on the possibility that video-games improve cognitive skills and that the benefits transfer to other domains (Green, Li, & Bavelier, 2010; Green, Pouget, & Bavelier, 2010). Just like with the video-game literature, a possibility that will have to be kept in mind in our meta-analysis is that the observed transfer from the source domains to the target domains might be due to confounds such as the placebo effect (Boot, Blakely, & Simons, 2011; Gobet et al., 2014).

Our meta-analysis[1] is an investigation of studies regarding the potential benefits of chess for children with respect to (a) mathematics skills, (b) reading skills, and (c) several cognitive skills (general intelligence, meta-cognition, attention/concentration, and spatial abilities). We chose these three categories of skills because they were the three categories chess-related research has been focusing on.

Our study had two main aims. The first aim was to estimate the overall effect size of the benefits of chess instruction by comparing experimental groups to control groups. The second aim was to evaluate the potential role of several factors in moderating the effect of chess instruction in children. The first four moderators addressed substantive aspects of the studies, and the last two covered methodological aspects:

1. Outcome: Mathematics, reading, or cognitive skills;
2. Duration of training (in hours);
3. Grade of the participants: Primary or secondary school;
4. Participants' category: Children with special educational needs or not;
5. Publication: Published or unpublished studies, where "published" is defined as having appeared in a peer-reviewed journal;
6. Design quality: Integer index (range 0–3, from poor to good) expressing the quality of the study design. The index measures three methodological characteristics: random allocation, administered pre-test, and avoidance of self-selection of the sample.

Along with the evaluation of the potential role of the above moderators, two specific sets of hypotheses were tested. The first pair of hypotheses dealt with the general question as to whether the skills acquired with chess instruction transfer to other domains. Two opposing hypotheses were tested. Hypothesis 1a predicted that, consistent with the literature on expertise and most of the literature on transfer, chess skill does not transfer to other domains, or at best the transfer is small and mostly due to unspecific factors (such as placebo effects). Hypothesis 1b, which reflects the view held by most researchers and practitioners in the field of chess instruction, predicted that there is substantial transfer. The second hypothesis dealt with the benefits of chess instruction on mathematics and reading. In line with Thorndike and Woodworth's (1901), it was predicted that transfer is stronger with mathematics than with reading, as chess shares more elements with the former topic than with the latter.

3. Method

3.1. Literature search

A systematic search strategy was used to find the relevant studies. The procedure is summarized in Fig. 1. Google Scholar, ProQuest (Dissertations & Theses), ERIC and Psyc-Info databases were searched to identify all the potential relevant studies. In addition, previous narrative reviews were examined, and we e-mailed researchers in the field asking for unpublished studies and data.

3.2. Inclusion/exclusion criteria

The studies were included according to the following seven criteria:

1. The design of the study was experimental or quasi-experimental; correlational and ex post facto studies were excluded.
2. The independent variable (chess instruction) was successfully isolated; the studies using chess instruction as one of several independent variables (such as other activities) in the experimental group were excluded.
3. The study presented a comparison between a chess intervention group and at least one control group.
4. The treatment and the control groups did not differ in terms of grade (e.g. third graders compared to fourth or fifth graders).
5. During the study, a measure of mathematical, reading, or cognitive skill was collected.
6. The participants of the study were pupils from kindergarten to the 12th grade.
7. The data presented in the published study were sufficient to calculate an effect size or the author(s) of the study, after having been contacted, provided the necessary data.

[1] Two previous meta-analyses were carried out on the effect of chess instruction: Benson (2006) and Nicotera and Stuit (2014). Neither calculated an overall effect size nor ran a moderator analysis. Rather, they divided the meta-analytic means into sub-categories (such as mathematics with chess instruction). The results they obtained were optimistic compared to ours, as they included several studies that were not included in the present meta-analysis because they did not satisfy the selection criteria.

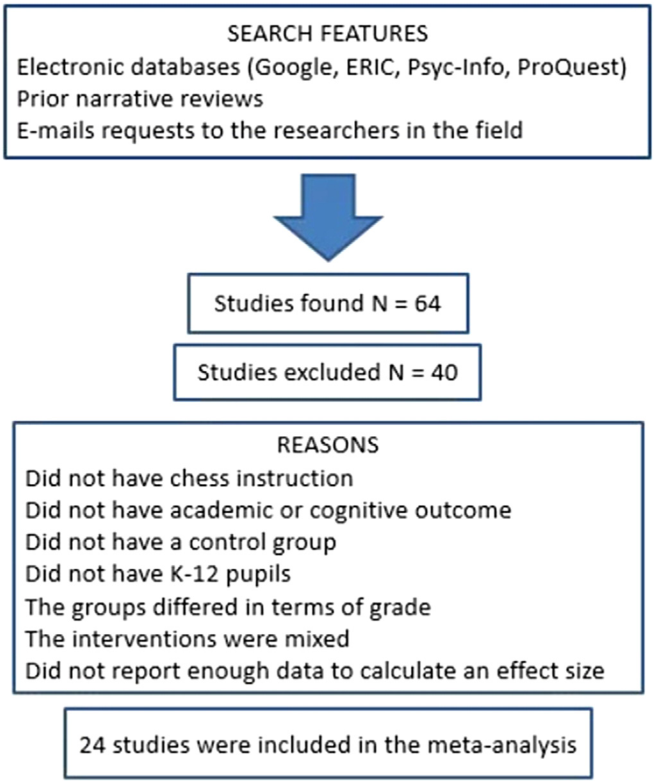

Fig. 1. Flow diagram of the studies considered and ultimately included in the meta-analysis.

We found 24 studies, conducted from July 1976 to July 2015, that met all the inclusion criteria (see Table 1). These studies included 25 independent samples and 40 effect sizes, and a total of 5221 participants (2788 in the experimental groups and 2433 in the control groups).

3.3. *Effect size*[2]

For the studies with an *only-post-test design*, the standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (Me - Mc)/SD_{pooled} \tag{1}$$

where SD_{pooled} is the pooled standard deviation and M_e and M_c are the means of the experimental group and the control group, respectively.[3] For the studies with a *repeated-measure design*, the standardized means difference was calculated with the following formula:

$$d = \left(M_{g-e} - M_{g-c}\right)/SD_{pooled-pre} \tag{2}$$

where $SD_{pooled\text{-}pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and of the control group, respectively. For the studies with an *ANCOVA design*, the standardized means difference was calculated with the following formula:

$$d = \left(M_{adj-e} - M_{adj-c}\right)\Big/SD_{pooled} \tag{3}$$

[2] All the formulas we used were taken from Schmidt and Hunter (2015).
[3] If the t or F statistics were provided, we used the regular formulas $d = t * ((n_1 + n_2)/(n_1 * n_2))$ and $d = (F * (n_1 + n_2)/(n_1 * n_2))$.

Table 1
Summary of the 24 studies included in the meta-analysis.

Study	Outcome	Published	Hours	Design quality	Special educational needs	Grade	Outcome measure
Aciego, Garcia, and Betancort (2012)	Cognitive	Yes	96	1	No	Both	WISC-R
Aydin (2015)	Maths & Cognitive	Yes	48	1	Yes	Secondary	Unknown
Barrett and Fish (2011)	Maths & Cognitive	Yes	25	2	Yes	Secondary	TAKS
Christiaen and Verhofstadt-Denève (1981)	Maths & Reading	Yes	42	2	No	Primary	DGB
DuCette (2009)	Maths & Reading	No	Not Given	0	No	Both	PSSA
Eberhard (2003)	Cognitive	No	60	1	Yes	Secondary	CogAT; NNAT
Forrest et al. (2005)	Maths & Reading	No	37	2	No	Primary	WISC-R (arithmetic subtest); Neale test
Fried & Ginsburg (n.d.)	Cognitive	No	Not Given	2	Yes	Primary	WISC-R
Garcia (2008)	Maths & Reading	No	90	1	No	Primary	TAKS
Gliga and Flesner (2014)	Cognitive	Yes	10	3	No	Primary	Krapelin test; Rey test
Hong and Bart (2007)	Cognitive	Yes	20	3	Yes	Both	RPM
Kazemi et al. (2012)	Maths & Cognitive	Yes	96	2	No	Both	TIMSS (mathematical literacy); Panaoura, Philippou & Christou test
Kramer & Filipp (n.d.)	Cognitive	No	32	2	No	Primary	Unknown
Margulies (1992)	Reading	No	Not Given	1	No	Primary	DRP
Rifner (1992)	Maths & Reading	No	30	2	No	Secondary	CTBS/4
Romano (2011)	Maths	No	25	3	No	Primary	INVALSI
Sala & Trinchero (in preparation)	Maths & Cognitive	No	10	3	No	Primary	OCDE-Pisa (mathematical literacy)
Sala et al. (2015)	Maths	Yes	18	3	No	Primary	OCDE-Pisa (mathematical literacy)
Sala, Gobet, Trinchero, & Ventura (submitted)	Maths & Cognitive	No	15	3	No	Primary	TIMSS (mathematical literacy); Panaoura & Philippou test
Scholz et al. (2008)	Maths & Cognitive	Yes	24	3	Yes	Primary	Arithmetic test designed by the authors; DL-KG
Sigirtmac (2012)	Cognitive	Yes	50	0	No	Primary	Unknown
Trinchero and Piscopo (2007)	Maths	No	30	2	No	Primary	Unknown
Trinchero & Sala (2016)	Maths	No	19	3	No	Primary	OCDE-Pisa (mathematical literacy)
Yap (2006)	Maths & Reading	No	50	0	No	Primary	Oregon State Assessment

where SD_{pooled} is the pooled standard deviation of the two standard deviations of the unadjusted means, and $M_{adj\text{-}e} - M_{adj-c}$ are the adjusted means of the experimental group and the control group, respectively. To correct for the upward bias, every Cohen's d was converted into Hedges's g by using the following formula:

$$g = d/(1 + 0.75/(N - 3)) \quad (4)$$

where N is the sample size of the study.Where reliability coefficients were available, the effect sizes were corrected for measurement error by using the following formula:

$$g' = g/a \quad (5)$$

where a is the square root of the reliability coefficient. It was possible to apply this correction to 31 effect sizes. Finally, there were three outliers whose residual errors had z scores greater than 4. These were Winsorized to z scores equal to 3.99.

Since we believed that the effect sizes had to reflect the actual improvement of the experimental groups and should not be the product of statistical artefacts, we adopted the following criterion: when the control group performance decreased in the post test, the effect size was calculated by considering $M_{g\text{-}c}$ (control group gain) equal to 0. Finally, the Comprehensive Meta Analysis (Version 3.0; Biostat, Englewood, NJ) software package was used for computing the effect sizes and conducting the statistical analyses.

4. Results

A random model ($K = 40$) was built to calculate an overall effect size.[4] The overall effect size was $g = 0.338$, 95% CI [0.242; 0.435], $p < .001$ (Fig. 2). The degree of heterogeneity between effect sizes was between moderate and high ($I^2 = 57.227$), suggesting the potential effect of some moderators. A trim-and-fill analysis showed that there was no publication bias.

[4] Twelve studies had more than one effect size. However, according to Tracz, Elmore, and Pohlmann (1992), violations of statistical independence have little or no effect on means, standard deviations, and confidence intervals.

Study name	Hedges's g	Lower limit	Upper limit	p-Value
	Statistics for each study			
Sala & Trinchero (in preparation) - M1	-0.114	-0.626	0.398	0.664
Gliga & Flesner (2014) - M1	-0.061	-0.840	0.718	0.877
Sala, Gobet, Trinchero, & Ventura (submitted) - M2	-0.030	-0.685	0.624	0.927
Eberhard (2003)	-0.028	-0.382	0.326	0.878
Scholz et al. (2008) - M1	0.020	-0.705	0.745	0.957
Romano (2011)	0.026	-0.078	0.130	0.624
Gliga & Flesner (2014) - M2	0.087	-0.669	0.843	0.822
Fried & Ginsburg (n.d.) - S2	0.100	-0.721	0.920	0.812
Forrest, Davidson, Stucksmith, & Glendinning (2005) - M1	0.101	-0.531	0.733	0.753
Aciego, Garcia, & Betancort (2012)	0.118	-0.185	0.422	0.444
Garcia (2008) - M1	0.122	-0.451	0.696	0.676
Scholz et al. (2008) - M2	0.122	-0.396	0.641	0.643
Fried & Ginsburg (n.d.) - S1	0.125	-0.698	0.947	0.766
Rifner (1992) - M2	0.147	-0.634	0.929	0.712
Yap (2006) - M2	0.152	-0.096	0.399	0.230
Hong & Bart (2007)	0.152	-0.473	0.777	0.633
Rifner (1992) - M1	0.173	-0.628	0.975	0.672
Forrest, Davidson, Stucksmith, & Glendinning (2005) - M2	0.236	-0.368	0.841	0.444
Kramer & Filipp (n.d.) - M2 *	0.262	-0.039	0.562	0.088
DuCette (2009) - M2	0.263	0.026	0.501	0.030
Yap (2006) - M1	0.273	0.025	0.520	0.031
Margulies (1992)	0.275	0.077	0.474	0.007
Christiaen & Verhofstadt-Denève (1981) - M1	0.280	-0.313	0.873	0.355
Sala & Trinchero (in preparation) - M2	0.287	-0.345	0.920	0.374
Sala, Gobet, Trinchero, & Ventura (submitted) - M1	0.333	-0.304	0.969	0.306
Trinchero & Sala (2016)	0.344	0.168	0.520	0.000
Garcia (2008) - M2	0.364	-0.221	0.948	0.222
DuCette (2009) - M1	0.373	0.138	0.608	0.002
Christiaen & Verhofstadt-Denève (1981) - M2	0.410	-0.186	1.006	0.177
Trinchero & Piscopo (2007)	0.411	0.027	0.795	0.036
Sala, Gorini, & Pravettoni (2015)	0.454	0.227	0.681	0.000
Gliga & Flesner (2014) - M3	0.563	-0.072	1.198	0.082
Kramer & Filipp (n.d.) - M1 *	0.627	0.319	0.936	0.000
Kazemi, Yektayar, & Abad (2012) - S2 - M2	0.649	0.272	1.027	0.001
Kazemi, Yektayar, & Abad (2012) - S2 - M1 *	0.743	0.378	1.108	0.000
Kazemi, Yektayar, & Abad (2012) - S1 - M2	0.790	0.245	1.335	0.004
Sigirtmac (2012) *	1.060	0.540	1.581	0.000
Kazemi, Yektayar, & Abad (2012) - S1 - M1 *	1.193	0.607	1.779	0.000
Barrett & Fish (2011)	1.232	0.457	2.007	0.002
Aydin (2015)	1.657	0.836	2.478	0.000
	0.338	0.242	0.435	0.000

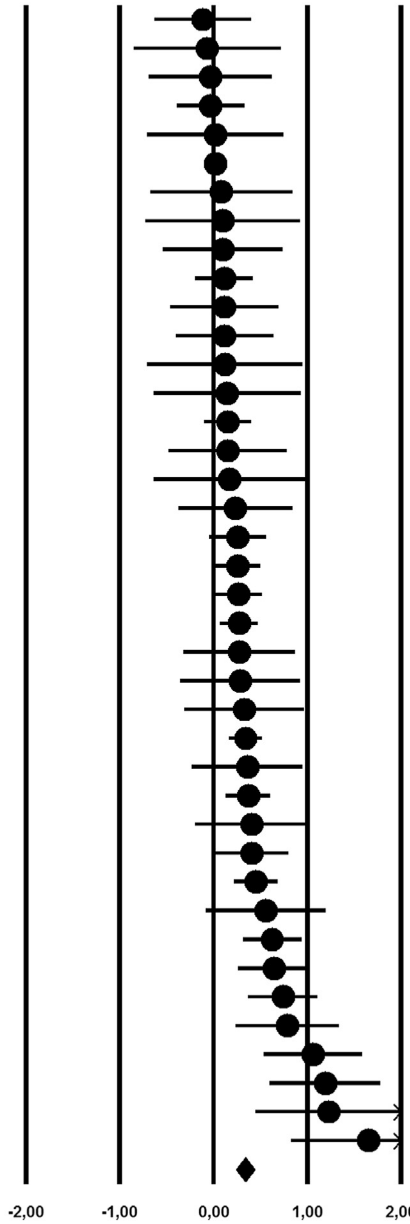

Fig. 2. Overall effect size (*g*) for chess training groups compared to control groups. Hedges's *g*s (circles) and 95% CIs (lines) are displayed for all effects entered into the meta-analysis. The diamond at bottom represents the meta-analytically weighted mean Hedges's *g*. For studies with multiple independent samples, the result of each sample (S1, S2, etc.) is reported separately. Analogously, for studies with multiple outcome measures, the result of each measure (M1, M2, etc.) is reported separately. Asterisks identify adjusted (Winsorized) outliers.

Consistent with this, a funnel plot analysis, depicting the relationship between standard error and effect size, was approximately symmetrical (see Fig. 3).

4.1. Moderator analyses

The only two statistically significant moderators were Duration of Training, which positively affected the effect sizes ($Q(1) = 3.89$, $b = 0.0038$, $p < .05$, two tailed, $K = 35$), and Publication, which also positively affected the effect sizes ($Q(1) = 10.17$, $b = 0.2941$, $p < .01$, two tailed, $K = 40$).

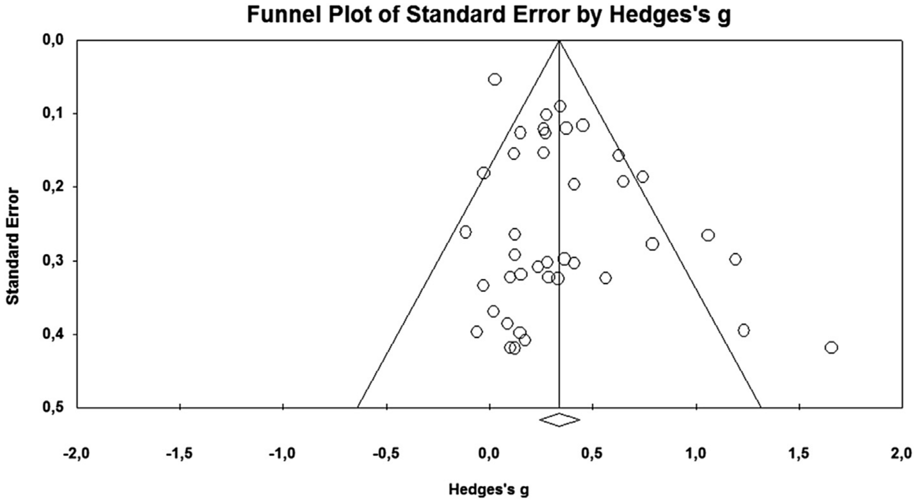

Fig. 3. Funnel plot of standard errors and effect sizes (*g*). The diamond at bottom represents the meta-analytically weighted mean Hedges's *g*.

Following Trinchero's (2012) suggestion (see Discussion), we considered 25 h as a threshold for the moderator Duration of Training. The overall effect size in studies with 25 or more hours of treatment was $g = 0.427$, 95% CI [0.271; 0.583], $p < .001$, $K = 23$, while the overall effect size in studies with less than 25 h of training was $g = 0.303$, 95% CI [0.189; 0.417], $p < .001$, $K = 12$. Regarding the moderator Publication, the overall effect size of the published studies was $g = 0.540$, 95% CI [0.346; 0.735], $p < .001$, $K = 17$, while the overall effect size of the unpublished studies was $g = 0.230$, 95% CI [0.149; 0.311], $p < .001$, $K = 23$.

4.2. Additional meta-analytic models

Although outcome was not a significant moderator, we ran three additional random models – one for each outcome – in order to investigate whether any outcome shows an overall effect size appreciably superior (or inferior, see discussion) to the others, as stated in Hypothesis 2.

The first model included the 17 mathematics-related effect sizes. The overall effect size was $g = 0.382$, 95% CI [0.229; 0.535], $p < .001$. A trim-and-fill analysis showed that there was no publication bias. The second model included the 16 cognitive-related effect sizes. The overall effect size was $g = 0.330$, 95% CI [0.130; 0.529], $p = .001$. A trim-and-fill analysis indicated that there was no publication bias. Finally, the third model included the seven reading-related effect sizes. The overall effect size was $g = 0.248$, 95% CI [0.128; 0.368], $p < .001$. A trim-and-fill analysis showed a possible publication bias (one study trimmed, left to the mean). The analysis showed that the point estimate was $g = 0.241$, 95% CI [0.122; 0.359].

5. Discussion

There is currently much research and excitement about the benefits of teaching chess in schools. The issue is theoretically important, since chess researchers' and practitioners' claims about the presence of far transfer are at variance with main theories of learning and expertise, which consider far transfer as difficult. In order to evaluate these diverging predictions, the current meta-analysis examined the effect exerted by chess instruction on academic (mathematics and reading) and cognitive abilities in children.

5.1. Substantive results

The first hypothesis predicted overall transfer beyond placebo effects. The results of the current meta-analysis suggest that chess instruction improves children's mathematical, reading, and cognitive skills moderately. Although this outcome seems promising, two considerations should be borne in mind. First, the overall effect size is not large enough to convincingly establish the effectiveness of chess instruction in enhancing the skills in consideration. By using Hattie's (2009) categorization, an overall effect size of $g = 0.338$ is not in the so-called "zone of desired effects," that is $d \geq 0.4$, which is the median value of the effectiveness of educational interventions estimated by Hattie's second-order meta-analysis. This suggests that chess instruction is no more effective in enhancing children's cognitive and academic skills than many (at least more than 50%) other possible educational interventions. Moreover, the observed difference between treatment and control groups might be due to chess instructors' passion rather than chess itself, because the potential role of placebo effects was rarely, if ever, controlled for in the studies under consideration (we will take up this methodological point below). Thus, consistent with Thorndike and Woodworth's (1901) common-element theory, the results tend to lend more support to Hypothesis 1a (chess skill does not transfer to other domains) than Hypothesis 1b (transfer will be substantial), which is largely held by the field of chess-in-school research. These considerations – along with the overall results of the meta-analysis – lead us to think that learning activities should be as close as possible to the skills to train; for example, mathematics instruction should be used to teach mathematical skills.

However, the positive influence of the hours of treatment on the results seems to support the idea that chess skill does transfer to other domains. Trinchero (2012) has suggested that appreciable positive effects occur only after 25–30 h of chess instruction. For studies with a minimum of 25 h of instruction, the overall g effect size was 0.427, which is a value in the "zone of desired effects" (see above). It is thus unlikely that this positive outcome is only the consequence of placebo effects, although this possibility cannot be ruled out completely. This suggests that 25–30 h of chess instruction is the minimum amount of instruction in order to obtain a significant transfer of learning from chess to other domains.

The second hypothesis, which was a more direct test of Thorndike and Woodworth's (1901) theory, predicted that transfer from chess should be stronger to mathematics than to reading, as chess shares more common elements with the former than the latter. Consistent with the hypothesis, the overall effect size was larger with mathematics than with reading ($g = 0.382$ vs. $g = 0.248$). Although outcome was not a significant moderator, reading seemed to benefit less from chess instruction than mathematics, as the effect size was substantially lower; this was despite the fact that five of the seven studies on reading used a long duration (30 h or more; no information about duration was available in the other two studies).

In the introduction, we presented Thorndike and Woodworth's (1901) view that transfer of skills occurs only between two domains that share components. It is plausible to argue that chess and mathematics have some components in common, such as their problem-solving nature and the importance of quantitative relationships. Therefore, the hypothesis that chess is a medium (in the sense of Feuerstein, Feuerstein, Falik, & Rand, 2006) through which cognitive skills are trained with some benefit for mathematics is plausible, even though it has not yet been convincingly supported by empirical research. However, with respect to reading, it is difficult to identify what components are shared with chess, unless we focus on very general commonalities (e.g., chess playing and reading are both decision-making activities). In their study of the effects of chess instruction on reading, Forrest, Davidson, Shucksmith, and Glendinning (2005) suggested that chess interventions enabled participants with low self-esteem to gain more confidence, which improved their literacy skills. If true, this suggestion – along with the small effect size ($g = 0.248$) – upholds the idea that the effects of chess interventions on reading are non-specific.

5.2. Methodological moderators

The index of design quality was not a significant moderator. This fact suggests that the results have not been significantly biased by the design used in the studies included in the meta-analysis. Nevertheless, as previously mentioned, the absence of an active control group in almost all the studies was a potential design-related confound we could not control for. The moderator Publication indicated that studies published in peer-reviewed journals have greater effect sizes. That studies with good results are more likely to be published is a common pattern in the literature (Schmidt & Hunter, 2015).

5.3. Limitations of this study

Regrettably, like the vast majority of studies carried out to assess the effect of educational methods, none of the studies considered in this review employed what Gobet and Campitelli (2006) called the "ideal design." This design includes the following requirements in addition to a treatment group: pre-test and post-test; two control groups (a do-nothing group and an active control group, necessary for removing the possibility of a placebo effect); random allocation to group; different personnel for conducting the pre-test, the treatment, and the post-test; and ideally – but nearly impossible to do in practice – experimenters' and testers' unawareness of the nature of the assignment into groups, and participants' unawareness of the goal of the experiment and the fact that they take part in an experiment. The presence of an active control group is crucial for controlling the possibility of placebo effects, and thus establishing the causal role of chess instruction in far transfer. Mechanisms that could produce "placebo effects" include instructors' motivation, the state of motivation induced by a novel activity, and educators' expectations (e.g., Boot, Simons, Stothart, & Stutts, 2013; Gobet & Campitelli, 2006). Without any active control group, it is not possible to exclude the possibility that positive results are due to such confounds, rather than to chess itself. It remains unknown whether a study with a more rigorous design would yield the same results as the studies previously conducted. Since nearly no study in the current meta-analysis had an active control group, which is necessary for ruling out possible placebo effects, the effects of chess instruction could have been systematically overestimated.

Another limitation of this field of research is that too few studies reliably controlled for moderator effects. In addition, the dependent variables were often very different between the studies: for example, basic arithmetic skills and mathematical problem-solving skills are not the same thing, and the same applies to meta-cognition, general intelligence, attention, and spatial abilities. We classified the studies using three broad kinds of outcomes (mathematical, reading, and cognitive skills) because, unfortunately, the small number of studies did not allow us to reliably evaluate the specific skills assessed as potential moderators.

5.4. Conclusions and recommendations for future research

Even if chess, under specific circumstances, seems to positively affect children's skills, there still are serious doubts about the real effectiveness of its practice. There is a need to clarify whether this positive influence is due to placebo effects or to chess instruction itself. In the latter case, research should identify the mechanisms underpinning the link between chess, the specific cognitive abilities involved and enhanced by the practice of the game, and their potential influence on mathematics and reading skills. In addition, the field should develop a detailed causal model explaining the cognitive processes that mediate learning and transfer. Finally, the data suggest that chess enhances children's mathematical skills and cognitive abilities more than reading skills, although the moderator analysis was not statistically significant. With reading skills, both the data and the explanations provided by researchers suggest that the positive effects of chess on children's reading skills are due to placebo effects. Further research should establish the reliability of these results.

Regarding future studies, we recommend to use an experimental design (random allocation, pre-tests and post-tests) with two control groups (a do-nothing group and an active control group). While logistically more complex, such a design is necessary in order to establish whether the benefits putatively provided by chess instruction are genuine and not caused by non-specific factors (e.g., placebo effect). Another important goal is to identify the specific characteristics of chess that might improve children's abilities, and which abilities they foster (e.g., attention, spatial abilities, quantitative reasoning, or metacognition). For example, is it the diversity of pieces on the board that help maintain attention? Does the movement of the pieces help to boost visuo-spatial abilities? Does chess ideally combine numerical, spatial, temporal, and combinatorial aspects? Does chess promote a better and more conscious way of thinking? In particular, it is important to demonstrate whether these features are common or not to other activities and games. Specifically, one should understand whether some features (e.g., quantitative relationships between pieces and problem-solving situations) are shared by other board games.

Thus, researchers should include (at least) two dependent variables – one academic and one cognitive – in their experimental designs, in order to shed some light on the causal relationships between chess instruction, and cognitive and academic skills. Many researchers, for instance, have claimed that chess enhances mathematical skills because chess practice relies on cognitive skills and mechanisms that, in turn, underlie mathematical skills. While this hypothesis is plausible, too few studies have directly addressed the question by assessing both a cognitive and an academic outcome, and the results have been contradictory. For example, Scholz et al. (2008) and Sala and Trinchero (in preparation) found no effect of chess on focused attention and meta-cognition respectively, whereas Kazemi, Yektayar, and Abad (2012) found a positive effect of chess practice on meta-cognitive abilities both in primary and in secondary school participants.

Finally, since the effectiveness of chess in enhancing children's intellectual skills seems to be dependent on the duration of the training, it would be useful to directly manipulate this variable in future studies, by systematically varying the duration of treatments between groups. This would ascertain the minimal and optimal amounts of chess instruction for far transfer: too short a duration might not provide enough time for progress, while too long a duration might lead to diminishing returns. Other worthwhile topics of investigation include a comparative study of different teaching methods with respect to their efficiency (e.g., is instruction better with computers or without computers? Are group activities preferable to individual activities, or is it the opposite? Are there more efficient orders of covering the material?). Finally, there has been little research that has explicitly mapped between chess and aspects of mathematics. Possible examples include bridging the chess board with the Cartesian graph and bridging the way the king moves in chess with block distance (as opposed to Euclidean distance). As it is known that awareness makes transfer more likely (Gick & Holyoak, 1980; Salomon & Perkins, 1989), it is plausible that making explicit the links between chess and mathematics could facilitate transfer.

In conclusion, the game of chess seems to exert a slight positive influence on both academic and cognitive abilities. Further research is needed to shed light on the relationship between cognitive and academic improvements, to evaluate the role of potential moderators and confounds, and to understand the role, if any, of placebo effects and game elements non-specific to chess.

References[5]

*Aciego, R., Garcia, L., & Betancort, M. (2012). The benefits of chess for the intellectual and socio-emotional enrichment in schoolchildren. *The Spanish Journal of Psychology, 15*, 551–559.

Anderson, J. R. (1990). *Cognitive psychology and its implication* (3rd ed.). New York: Freeman.

*Aydin, M. (2015). Examining the impact of chess instruction for the visual impairment on mathematics. *Educational Research and Reviews, 10*, 907–911.

*Barrett, D. C., & Fish, W. W. (2011). Our move. Using chess to improve math achievement for students who receive special education services. *International Journal of Special Education, 26*, 181–193.

Bart, W. M. (2014). On the effect of chess training on scholastic achievement. *Frontiers in Psychology, 5*, 762.

Benson, C. (2006). How training chess affects children's scores on academic aptitude tests: A meta-analysis. In T. Redman (Ed.), *Chess and education. Selected essays from the Koltanowski Conference* (pp. 144–160). Dallas, TX: University of Texas at Dallas.

Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence, 35*, 457–470.

Binev, S., Attard-Montalto, J., Deva, N., Mauro, M., & Takkula, H. (2011). *Declaration of the European Parliament, 0050/2011*.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology, 2*, 226.

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445–454.

[5] References marked with an asterisk indicate studies included in the meta-analysis.

Bühren, C., & Frank, B. (2010). *Chess players' performance beyond 64 squares: A case study on the limitations of cognitive abilities transfer*. MAGKS Joint Discussion Paper Series in Economics, 19–2010.
Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Hillsdale, NJ: Erlbaum.
*Christiaen, J., & Verhofstadt-Denève, L. (1981). Schaken en cognitieve ontwikkeling [[Chess and cognitive development]]. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 36*, 561–582.
Costa, A., & Kallick, B. (2009). *Learning and leading with habits of mind: 16 essential characteristics for success*. Alexandria (VA): Association for Supervision & Curriculum Deve.
De Corte, E., & Verschaffel, L. (1986). Effects of computer experience on children's thinking skills. *Journal for Structural Learning, 9*, 161–174.
Doll, J., & Mayr, U. (1987). Intelligenz und Schachleistung: Eine Untersuchung an Schachexperten [[Intelligence and performance in chess: a study of chess expert]]. *Psychologische Beitrage, 29*, 270–289.
Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (1999). *How people learn: Bridging research and practice*. Washington, DC: National Academies Press.
*DuCette, J. (2009). *An evaluation of the chess Challenge program of ASAP/after school activities Partnerships*. Philadelphia, PA: After School Activities Partnerships.
*Eberhard, J. W. (2003). *The relationship between chess instruction and verbal, quantitative, and nonverbal reasoning abilities of economically disadvantage students*. Corpus Christi, TX: Texas A&M University (Unpublished doctoral dissertation).
Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725–747.
Feuerstein, R., Feuerstein, R. S., Falik, L., & Rand, Y. (2006). *The Feuerstein instrumental enrichment program*. Jerusalem: ICELP Publications.
Forrest, D., Davidson, I., Shucksmith, J., & Glendinning, T. (2005). *Chess development in Aberdeen's primary schools: A study of literacy and social capital*. Aberdeen, Scotland: University of Aberdeen. Retrieved from www.gov.scot/Resource/Doc/930/0009711.pdf.
*Fried, S. and Ginzburg, N. (n.d.), The effect of learning to play chess on cognitive, perceptual and emotional development in children. Unpublished manuscript.
Frydman, M., & Lynn, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology, 83*, 233–235.
*Garcia, N. V. (2008). *Scholastic chess club participation and the academic achievement of Hispanic fifth grade students in south Texas*. Houston, TX: University of Houston (Unpublished doctoral dissertation).
Garner, R.. (2012). Chess makes a dramatic comeback in primary schools. http://www.independent.co.uk/news/education/education-news/chess-makes-a-dramatic-comeback-in-primary-schools-8301313.html
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.
*Gliga, F., & Flesner, P. I. (2014). Cognitive benefits of chess training in novice children. *Procedia – Social and Behavioral Sciences, 116*, 962–967.
Gobet, F. (2011). *Psychologie du talent et de l'expertise [Psychology of talent and expertise]*. Bruxelles: De Boeck.
Gobet, F. (2015). *Understanding expertise: A multi-disciplinary approach*. London: Palgrave/Macmillan.
Gobet, F., & Campitelli, G. (2002). Intelligence and chess. In J. Retschitzki, & R. Haddad-Zubel (Eds.), *Step by step. Proceedings of the 4th colloquium board games in academia* (pp. 103–112). Fribourg (Switzerland): Editions Universitaires.
Gobet, F., & Campitelli, G. (2006). Educational benefits of chess instruction. A critical review. In T. Redman (Ed.), *Chess and education. Selected essays from the Koltanowski Conference* (pp. 124–143). Dallas, TX: University of Texas at Dallas.
Gobet, F., de Voogt, A. J., & Retschitzki, J. (2004). *Moves in mind*. Hove, UK: Psychology Press.
Gobet, F., Johnston, S. J., Ferrufino, G., Johnston, M., Jones, M. B., Molyneux, A., & Weeden, L. (2014). No Level Up!': no effects of video game specialization and expertise on cognitive performance. *Frontiers in Psychology, 28*(5), 1337.
Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: a psychometric investigation. *Acta Psychologica, 124*, 398–420.
Green, C. S., Li, R. J., & Bavelier, D. (2010a). Perceptual learning during action video game playing. *Topics in Cognitive Science, 2*, 202–216.
Green, C. S., Pouget, A., & Bavelier, D. (2010b). Improved probabilistic inference as a general learning mechanism with action video games. *Current Biology, 20*, 1573–1579.
Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy, 24*, 23–37.
Gurtner, J. L., Gex, C., Gobet, F., Nunez, R., & Restchitzki, J. (1990). La récursivite rend-elle l'intelligence artificielle? [Does recursion make intelligence artificial?]. *Revue Suisse de Psychologie, 49*, 17–26.
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–51.
Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2012). *Achievement growth: international and US state trends in student performance*. Cambridge, MA: Program on Education Policy and Governance, Harvard Kennedy School.
Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
*Hong, S., & Bart, W. M. (2007). Cognitive effects of chess instruction on students at risk for academic failure. *International Journal of Special Education, 22*, 89–96.
Kazemi, F., Yektayar, M., & Abad, A. M. B. (2012). Investigation the impact of chess play on developing meta-cognitive ability and math problem-solving power of students at different levels of education. *Procedia – Social and Behavioral Sciences, 32*, 372–379.
*Kramer, A. & Filipp, S. (n.d.). Chess at Tryer-Olewig Primary School: Summary and the evaluation of the outcomes of the German School Chess Foundation (short version). Unpublished manuscript.
Margulies, S. (1992). *The effect of chess on reading scores: District Nine chess program; Second year report*. New York, NY: The American Chess Foundation.
Mestre, J. P. (2005). *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age.
Nicotera, A., & Stuit, D. (2014). *Literature review of chess studies*. Denver, CO: Basis Policy Research.
Perkins, D. N., & Grotzer, T. A. (2000). Teaching intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 492–515). Cambridge: Cambridge University Press.
Perkins, D. N., & Salomon, G. (1994). Transfer of learning. In T. N. Postlethwaite, & T. Husen (Eds.), *International encyclopedia of education* (pp. 6452–6457). Oxford: Elsevier.
Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist, 47*, 189–203.
*Rifner, P. J. (1992). *Playing chess: A study of the transfer of problem-solving skills in students with average and above average intelligence*. West Lafayette, IN: Purdue University (Unpublished doctoral dissertation).
Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (2002). On the constraints of encapsulated knowledge: clinical case representations by medical experts and subexperts. *Cognition and Instruction, 20*, 27–45.
*Romano, B. (2011). *Does playing chess improve math learning? Promising (and inexpensive) results from Italy*. Philadelphia, PA: University of Pennsylvania (Unpublished doctoral dissertation).
*Sala, G., Gobet, F., Trinchero, R., & Ventura, S. (2016). *Does chess instruction enhance mathematical ability in children? A three-group design to control for placebo effects*. submitted for publication (Unpublished manuscript).
Sala, G., Gorini, A., & Pravettoni, G. (2015). Mathematical problem solving abilities and chess: An experimental study on young pupils. *SAGE Open, July-September*, 1–9.
Sala, G., & Trinchero, R. (2016). *Is meta-cognition the link between chess training and improvement in mathematics? A study on primary school children* (in preparation) (Unpublished manuscript).
Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 113–142.
Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Newbury Park, CA: Sage.

Schneider, W., Gruber, H., Gold, A., & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimenal Child Psychology, 56*, 328–349.
Scholz, M., Niesch, H., Steffen, O., Ernst, B., Loeffler, M., Witruk, E., et al. (2008). Impact of chess training on mathematics performance and concentration ability of children with learning disabilities. *International Journal of Special Education, 23*, 138–148.
*Sigirtmac, A. D. (2012). Does chess training affect conceptual development of six-year-old children in Turkey? *Early Child Development and Care, 182*, 797–806.
Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 9*, 374–382.
Tracz, S. M., Elmore, P. B., & Pohlmann, J. T. (1992). Correlational meta-analysis: Independent and nonindependent cases. *Educational and Psychological Measurement, 52*, 879–888.
Trinchero, R. (2012). *Gli scacchi, un gioco per crescere. Sei anni di sperimentazione nella scuola primaria [Chess, a game to grow up with. Six year of research in primary school]*. Milan: Franco Angeli.
*Trinchero, R., & Piscopo, M. (2007). *Gli scacchi, un gioco per crescere. [Chess, a game to grow up with]* (Unpublished manuscript).
Trinchero, R., & Sala, G. (2016). Can chess training improve Pisa scores in Mathematics? an experiment in Italian primary schools. *Eurasia Journal of Mathematics, Science & Technology Education, 12*, 655–668.
Unterrainer, J. M., Kaller, C. P., Leonhart, R., & Rahm, B. (2011). Revising superior planning performance in chess players: the impact of time restriction and motivation aspects. *American Journal of Psychology, 124*, 213–225.
Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuospatial abilities in chess players. *British Journal of Psychology, 30*, 303–311.
*Yap, K. O. (2006). *Chess for success evaluation: Final report*. Portland, OR: Northwest Regional Educational Laboratory.