Recent theorizing on second language (L2) motivation has proposed viewing motivation as a function of the language learners’ vision of their desired future language selves. This would suggest that the intensity of motivation is partly dependent on the learners’ capability to generate mental imagery. In order to test this hypothesis, this study investigates whether learner characteristics are related to sensory and imagery aspects with indices of the strength of the learners’ future L2 self-guides (ideal and ought-to L2 selves) and how these variables are linked to learning achievement in two target languages, English and Mandarin, assessed both by self-report and objective measures. One hundred seventy-two Year 8 Chinese students (ages 13–15) completed a questionnaire survey, and the results reveal several significant associations between the future self-guides and intended effort and actual grades, including a consistently positive relationship between the ideal self and the criterion measures. The findings also confirm the multisensory dimension of future self-guides, suggesting the importance of a broad imagery capacity (including both visual and auditory components) in the development of individuals’ future self-identities. Finally, the ideal-self images associated with different languages were shown to form distinct L2-specific visions, which has various implications for future research with regard to the potential positive or negative interaction of these self images.

**Keywords**  L2 motivational self system; ideal L2 self; ought-to L2 self; imagery; sensory styles
Introduction

The field of second language (L2) motivation research has had more than 50 years of continuous history and has witnessed various phases of development, with important milestones including Gardner and Lambert’s (1959, 1972) pioneering research that introduced the concepts of integrative and instrumental orientation, Clément’s (1980) addition of self-confidence to the motivational research paradigm, Gardner’s (1985) and Gardner and MacIntyre’s (1992, 1993) detailed discussion of the socioeducational model in second language acquisition (SLA), and the educational shift in motivation research in the 1990s (e.g., Crookes & Schmidt, 1991; Dörnyei, 1994). The latest emerging theoretical paradigm has been the L2 Motivational Self System proposed by Dörnyei (2005, 2009), which focuses on L2 learners’ self-perception, particularly the perception of their desired future self-states. This paradigm has drawn on Possible Selves Theory (Markus & Nurius, 1986, 1987) and Self-Discrepancy Theory (Higgins, 1987; Higgins, Klein, & Strauman, 1985), which suggest that possible selves—that is, individuals’ ideas of what they might become, what they would like to become, and what they are afraid of becoming in the future—can influence behavior by highlighting the discrepancies between the current actual and the future desired selves. According to the theory, this noticed discrepancy will generate affect which will in turn be translated into action to reduce the gap. In order to apply Possible Selves Theory to the domain of SLA, Dörnyei has proposed a parsimonious motivation construct made up of three constituents:

1. **Ideal L2 Self**, which concerns the L2-specific facet of one’s ideal self: if the person we would like to become speaks an L2 (e.g., the person we would like to become is associated with traveling or doing business internationally), the ideal L2 self is a powerful motivator to learn the L2 because we would like to reduce the discrepancy between our actual and ideal selves.

2. **Ought-to L2 Self**, which concerns the attributes that individuals believe they ought to possess to avoid possible negative outcomes; such perceived duties, external expectations, and obligations may therefore bear little resemblance to the individual’s own desires or wishes.

3. **L2 Learning Experience**, which concerns situation-specific motives related to the immediate learning environment and experience (e.g., the positive impact of success or the enjoyable quality of a language course).
Thus, the L2 Motivational Self System suggests that there are three primary sources of the motivation to learn an L2: (a) the learners’ internal desire to become an effective L2 user, (b) social pressures coming from the learner’s environment to master the L2, and (c) the actual experience of being engaged in the L2 learning process.

Since the development of Dörnyei’s (2005) tripartite theory, a number of studies have been conducted using the L2 Motivational Self System as their theoretical basis (e.g., Csizér & Lukács, 2010; Dörnyei & Ushioda, 2009a; Henry 2009, 2010, 2011; Hiver, in press; Kormos, Kiddle, & Csizér, 2011; Magid, 2012; Papi, 2010). The findings have typically confirmed the overall explanatory power of the model, with the ideal L2 self, in particular, seen as a strong predictor of various criterion measures related to language learning, playing a substantive role in determining motivated behavior. With regard to specific issues, research has shown that the strength of the ideal L2 self increased significantly in teenage girls but decreased in boys over a period of 4 years in Sweden, resulting in a “gender role intensification” (Henry, 2009). In Iran, Papi (2010) found that, while the ideal L2 self and the L2 learning experience decreased students’ English anxiety, the ought-to L2 self significantly made them more anxious. An interesting line of research into third language (L3) attainment revealed that a strong English-related ideal self may act as a source of interference when learning another foreign language (Henry, 2010; Cszigér & Lukács, 2010). On the other hand, some studies have found only a marginal relationship between the ought-to L2 self and motivated L2 behavior (Cszigér & Kormos, 2009; Cszigér & Lukács, 2010), and Kim (2009a), for example, suggests that learners need to internalize their ought-to L2 self sufficiently for it to exert its motivational influence.

The present study investigates the claim that the intensity of motivation is partly dependent on the learners’ capability to generate mental imagery. Specifically, focusing on a sample of 13- to 15-year-old Chinese-background learners in two target languages, English and Mandarin, and collecting self-reported evidence as well as objective measures, we sought to illuminate the relationship among learner characteristics related to sensory and imagery aspects, indices of the strength of the learners’ future L2 self-guides (ideal and ought-to L2 selves), and learning achievement.

Background to the Study

Vision

A key aspect of the L2 Motivational Self System is imagery, a phenomenon that has been highlighted in various areas of L2 learning in the past few decades,
for example, in grammar teaching (Gerngross, Puchta, & Thornbury, 2006), vocabulary learning (Cohen, 1987; Ellis & Beaton, 1993; Shen, 2010; Stevick, 1986), reading (Arnold, 1999; Green & Donahue, 2009; Krasny & Sadoski, 2008), writing (Wright & Hill, 2008), and listening comprehension (Center, Freeman, Robertson, & Outhred, 1999). These studies have accentuated the various ways of using images and imagination to empower L2 learners in acquiring an L2. Our present focus is on the role images and senses play in shaping the motivation to learn an L2 through promoting a more vivid mental representation of one’s self in future states. This interest goes back to Markus and Nurius’s (1986) original conceptualization of possible selves, as these scholars emphasized in their seminal paper that possible selves are represented in the mind in the same imaginary and semantic way as the here-and-now self; that is, they are a reality for the individual: People can see and hear their possible future self (see also Ruvolo & Markus, 1992). This means that, in many ways, possible selves are similar to one’s visions about oneself; indeed, Markus and Nurius (1987, p. 159) confirm that “[p]ossible selves encompass within their scope visions of desired and undesired end states” (emphasis added)—thus, possible selves can be seen as the “vision of what might be” (Markus & Nurius, 1987, p. 159).

A vision can be defined as “a mental representation that occurs without the need for external sensory input” (Stopa, 2009, p. 1). Although these mental images involve various forms of imaginary perception (visual, auditory, olfactory, or tactile) conjured up deliberately for a particular purpose, research has found that mental imagery relies to a large extent on the same neural mechanisms and pathways as actual perception, and studies of brain damage have also shown that such injuries often produce parallel deficits in imagery and in perception (for reviews, see, e.g., Moulton & Kosslyn, 2009; Reisberg & Heuer, 2005). This stimulatory nature of mental imagery is at the heart of its potency, and in the context of L2 learning, Dörnyei (2009) has also highlighted the central role of images and senses as integral components of desired future self-images, stating that it is the “experiential element that makes possible selves larger than any combinations of goal-related constructs” (p. 15). The key assertion is that learners with a vivid and detailed ideal self-image that has a substantial L2 component are more likely to be motivated to take action in pursuing language studies than their peers who have not articulated a desired future goal-state for themselves (for an overview, see Dörnyei & Kubanyiova, in press).

While the motivational function of visualization has been an established fact in sports psychology since Paivio’s (1985) seminal paper on the role of imagery in sport performance (see also Hall, Mack, Paivio, & Hausenblas,
there has been rather limited research in this area within the L2 literature. Some validation of the imagery–motivation link has been offered by recent intervention studies in which various possible selves enhancement activities were employed to facilitate future identity formation and to strengthen students’ future self images (Fukada, Fukuda, Falout, & Murphey, 2011; Jones, 2012; Magid & Chan, 2012; Sampson, 2012). These have reported that most participants have found visualization tasks focusing on their future self-guides motivating and they tended to invest more effort in language learning as a result of the program, thereby attesting to the impact of the treatment.

**Imagery Capacity and Preferences**

Besides the intervention studies mentioned above, the connection between imagery skills and motivation has been examined by another series of empirical studies over the past 5 years that specifically examined the motivational relevance of imagination and sensory preferences (e.g., visual style), and this research program is central to the current study. First, a pilot study by Al-Shehri (2009), based on his master’s research conducted at Nottingham University, yielded some promising results in this respect. The researcher investigated the relationship of L2 learners’ visual learning style preferences and self-reported imaginative capacities with their motivation in a sample of 200 Saudi learners of English (98 university and 102 secondary school students) to test the hypothesis that learners who exhibit a visual learning style preference are more likely to possess stronger capacity for visual imagery and imagination and are therefore more likely to develop a stronger ideal L2 self than visually less capable peers. In accordance with this hypothesis, the obtained results revealed strong, significant positive associations between students’ visual styles, imagination, ideal L2 selves, and motivated L2 behavior. Visual style was found to be highly correlated with imagination, the ideal L2 self, and motivated behavior, indicating a robust interrelationship among these factors.

Inspired by these results, Kim and his colleague conducted two follow-up studies in Korea, also incorporating other perceptual learning preferences such as auditory and kinesthetic styles. Kim’s (2009b) survey of 974 Korean primary school students confirmed the positive association between motivation and visual style and also found a significant positive correlation between auditory learning style and imagination, the ideal L2 self, and motivated behavior. Kim and Kim (2011) extended the inquiry from primary to secondary school students and found in a sample of 495 Korean high school learners significant correlations between their ideal English self and a number of factors, including
visual and auditory learning style preferences and imagery capacity, suggesting that these are key components in the formation of a vivid ideal L2 self.

Finally, in her master’s research at Nottingham University, Eid (2008) examined the cross-linguistic aspects of imagery by investigating 93 Cypriot secondary school learners of three target languages, English, Italian, and French. As can be seen in Table 1, which presents Eid’s main findings alongside the results of Al-Shehri (2009), Kim (2009b), and Kim and Kim (2011), the ideal L2 self was found to explain a significant amount of variance of achievement in all three target languages, whereas the sensory and imagery variables produced mixed results. However, because of the nature of her investigation (a master’s study with limited resources and time), no data were collected about how the three languages related to each other in terms of subjective learner preference and objective situational affordances such as time allocation and interethnic contact opportunities. Yet, this and the other studies reported above have produced sufficient evidence to make us realize that imagery matters, thereby warranting further systematic studies.

**The Operationalization of Imagery Capacity and Sensory Preferences**

How can we operationalize imagery capacity and sensory preferences for research purposes? The term “imagery capacity” in this article bears the same

---

**Table 1 Correlations between the ideal L2 self and various factors as identified in four studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ideal English self</td>
<td>Ideal English self</td>
<td>Ideal English self</td>
<td>Ideal French self</td>
</tr>
<tr>
<td>Imagery capacity</td>
<td>.46**</td>
<td>.26**</td>
<td>.27**</td>
<td>.30**</td>
</tr>
<tr>
<td>Visual style</td>
<td>.65**</td>
<td>.34**</td>
<td>.39**</td>
<td>-.00</td>
</tr>
<tr>
<td>Auditory style</td>
<td>-</td>
<td>.29**</td>
<td>.21**</td>
<td>-.02</td>
</tr>
<tr>
<td>Kinesthetic style</td>
<td>-</td>
<td>.03</td>
<td>-.14**</td>
<td>-</td>
</tr>
<tr>
<td>Motivated behavior</td>
<td>.78**</td>
<td>.58**</td>
<td>.72**</td>
<td>.45**</td>
</tr>
<tr>
<td>Grades</td>
<td>-</td>
<td>-</td>
<td>.23**</td>
<td>.17</td>
</tr>
</tbody>
</table>

*Note. A missing coefficient means that the variable was not assessed in the particular study. **p < .01.*
meaning as the concept of “imagery ability” in sports psychology, where it is defined as “an individual’s capability of forming vivid, controllable images and retaining them for sufficient time to effect the desired imagery rehearsal” (Morris, 1997, p. 37). It encompasses a range of sensory modalities (e.g., visual, auditory) and several dimensions such as vividness, relative ease of control, and duration of an image (Morris, Spittle, & Watt, 2005). While there exists a variety of approaches in imagery assessments, the most common assessment format involves the use of self-report questionnaires (Sheehan, Ashton, & White, 1983; for a detailed discussion, see Richardson, 1994).

There are no established templates for the measurement of imagery capacity in L2 learning; therefore previous studies have followed two approaches. On the one hand, they developed self-report items (typically in a Likert scale format) drawing on examples from the general psychological literature (as summarized by Richardson, 1994). The “imagery capacity” scale in the current study falls into this category. These scales, then, need to be validated through using the standard procedures of questionnaire design (Dörnyei, 2010), which typically involves a combination of content analysis and quantitative item analysis by means of computing Cronbach’s alpha internal consistency reliability coefficients.

On the other hand, previous studies have also drawn on established L2 learning style inventories (Cohen, Oxford, & Chi, 2001; Oxford, 1993; Reid, 1995) and adapted items assessing visual and auditory styles for their purposes. We have also adopted this second option based on the following consideration. While in the learning styles literature visual and auditory style preferences have typically been used to place students on a visual–auditory continuum (Dörnyei, 2005), the actual measurement of these styles usually involves separate numeric rating scales for both the visual and the audio components (rather than a comparison or a forced choice between them). Therefore, these scales are measured by graded response options (e.g., marking one’s response on a 1–5 scale), thereby not only indicating preference but also strength (e.g., marking “5” indicates a stronger relevance than marking “3”). Consequently, a high score on these scales indicates, in effect, highly developed sensory processing skills in L2 learning as reported by the student. This explains the common observation in the past that learners can be equally high or low in both style dimensions rather than displaying a marked preference for one or the other.

There is one final issue to address concerning sensory preferences: While it may be logical to include the assessment of the learners’ visual style in the study of vision, why would auditory style be relevant in this respect? We have seen in the introduction that mental images can involve various forms
of perception (visual, auditory, olfactory, or tactile) and, indeed, experiencing ourselves as an agent in a future situation—which is what future self-guides are about—does logically encompass various sensual dimensions of the experience (cf. Erikson, 2007). For example, athletes report that they can hear the crowd roaring in the stadium on the imagined day of the crucial competition; it is also a fairly common experience for many people to imagine conversations in which they can hear themselves and the interlocutor’s responses. Findings in psychology also affirm that nonvisual imageries are comparable to visualized ones; for example, Eardley and Pring (2006) found that auditory cue words (e.g., thunder, bark) were just as effective as visual cue words (e.g., sunset, moon) in simulating future events. Thus, we are in agreement with Knäuper, Roseman, Johnson, and Krantz’s (2009) conclusion that, in mental imagery, “individuals mentally mimic perceptual, motor, and emotional experiences, resulting in mental representations of the imagined objects, situations, emotions and actions in multiple sensory modalities including, but not limited to, vision” (p. 182).

Research Focus

Because of the potential theoretical relevance of the visionary aspect of L2 motivation and the intriguing results reported in previous studies, we set out to contribute to this line of inquiry by drawing together the main variables studied in the past in a single investigation, while also adding further components to the research design. In order to investigate multiple target languages, we selected a learning environment (Hong Kong) where two prestigious languages—English and Mandarin—are learned in a parallel manner, and we collected both self-report questionnaire data and objective course achievement grades to serve as criterion measures. We also extended the scope of the research paradigm by including the assessment of the ought-to L2 self to see how vision that is imposed on the learner externally compares to the learners’ self-generated ideal self imagery. This inclusion might be particularly relevant in a Chinese context where, as Magid (2012) and Shek and Chan (1999) explain, family duties and obligations play a more important role in shaping student motivation than in many Western learning contexts. Based on theoretical considerations, we formed the following three research hypotheses:

H1: In accordance with previous findings, both the ideal and the ought-to L2 selves will be positively associated with the two criterion measures, intended effort and course grades, for both English and Mandarin.
H2: There will be positive correlations between the future self-guides and the (a) visualization-related variables and (b) auditory-style variables.

(Because earlier research did not find any meaningful relationship between future self-guides and kinesthetic style, this variable was not included in the research paradigm.)

H3: Because the association of imagery with motivation is assumed to concern the very foundation of vision (as a technical term defined earlier), this association will be L2-independent, that is, not influenced by the nature of the target language involved.

Besides our primary focus on the imagery dimension of motivation, the inclusion of multiple target languages also allows for the examination of a fundamental question related to future self-guides, raised by Dörnyei and Ushioda (2009b), namely, how distinct the self-images associated with different target languages are. In other words, do learners have several different desired possible self images of themselves, as Markus and Nurius (1986) assume, or one broad ideal and one ought-to self construct with various facets as Higgins (1987) proposes? We did not have any specific hypotheses addressing this question because past theorizing has produced arguments potentially pointing to both options.

**Method**

**Participants**
The sample consisted of 172 Year 8 students in Hong Kong, 82 boys and 88 girls (2 with missing gender data), 13 to 15 years of age, who were all of Chinese ethnicity and spoke Cantonese as their first language. They were studying both English and Mandarin at a lower intermediate level in a Band 1 secondary school (which denotes the highest intake of successful students on a three-point scale: Bands 1–3). Within the Hong Kong learning context both languages have obvious relevance and ethnolinguistic vitality for the learners, and the value of Mandarin instruction has recently increased even more with the initiation of the Individual Visit Scheme, which allows independent travelers from Mainland China to visit Hong Kong, making these visitors a salient and commercially highly profitable group to interact with.
Instrument

The instrument used in this study involved a self-report questionnaire focusing on the two target languages, English and Mandarin. The motivational measures were based on Taguchi, Magid, and Papi’s (2009) questionnaire, the visual and auditory style scales on Cohen et al.’s (2001) Learning Style Survey (LSS) and Reid’s (1984) Perceptual Learning Style Preference Questionnaires (PLSPQ), and the imagery capacity scale on established imagery measures in Richardson (1994). The criterion variables measured in our survey included measures of intended learning effort and actual achievement grades in both languages. The self-report items were assessed using 5-point Likert-type scales, ranging from strongly disagree (1) to strongly agree (5). The main variable groups in the questionnaire were as follows:

1. **Motivation**: ideal L2 self and ought-to L2 self (five items each for both English and Mandarin);
2. **Sensory styles**: visual learning style (five items) and auditory learning style (five items);
3. **Imagery capacity**, that is, the ability to create visual imagery in one’s mind (five items; see the Appendix);
4. **Criterion measures**: self-reported intended learning effort (five items for both English and Mandarin) and actual achievement grades in end-of-term L2 exams.

The English version of the questionnaire was translated into Cantonese by the second author and was then edited by a professional English–Cantonese translator. Before administering it to the students in this particular school, the questionnaire was piloted in two classes \( n = 60 \) in two independent secondary schools in Hong Kong to explore whether students had any difficulties comprehending the questions. The administrators were asked to take note of any problems encountered by the students but no comprehension issues were recorded in the two classes.

Data Collection and Analysis

The survey was administered to five Year 8 classes by the same English teacher during home period in which students were informed about the study and the purposes prior to the actual survey. The students took approximately 15 to 20 minutes to complete the questionnaire. The data were coded in SPSS version 18.0. After computing Cronbach’s alpha internal consistency reliability coefficients, five inferential statistical procedures were employed: correlations and multiple correlations to establish relationships among variables, exploratory
Table 2 Information about the multi-item scales

<table>
<thead>
<tr>
<th>Variables</th>
<th>No. of items</th>
<th>Cronbach’s alpha</th>
<th>Sample item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal English self</td>
<td>5</td>
<td>.78</td>
<td>I often imagine myself speaking these languages as if I were a native speaker of these languages.</td>
</tr>
<tr>
<td>Ideal Mandarin self</td>
<td>5</td>
<td>.81</td>
<td>I often imagine myself speaking these languages as if I were a native speaker of these languages.</td>
</tr>
<tr>
<td>Ought-to English self</td>
<td>5</td>
<td>.77</td>
<td>I have to study these languages, because, otherwise, I think my parents will be disappointed with me.</td>
</tr>
<tr>
<td>Ought-to Mandarin self</td>
<td>5</td>
<td>.76</td>
<td>I have to study these languages, because, otherwise, I think my parents will be disappointed with me.</td>
</tr>
<tr>
<td>Intended effort in English</td>
<td>5</td>
<td>.77</td>
<td>I am prepared to expend a lot of effort in learning these languages.</td>
</tr>
<tr>
<td>Intended effort in Mandarin</td>
<td>5</td>
<td>.80</td>
<td>I am prepared to expend a lot of effort in learning these languages.</td>
</tr>
<tr>
<td>Visual style</td>
<td>3</td>
<td>.49</td>
<td>I learn more by reading textbooks than by listening to lectures.</td>
</tr>
<tr>
<td>Auditory style</td>
<td>5</td>
<td>.69</td>
<td>I learn better in class when the teacher gives a lecture.</td>
</tr>
<tr>
<td>Imagery capacity</td>
<td>5</td>
<td>.68</td>
<td>If I wish, I can imagine some things so vividly that they hold my attention as a good movie or story does.</td>
</tr>
</tbody>
</table>

factor analyses to identify factors that represent the underlying relationships across the two target languages, paired-samples t-tests to compare the magnitude of self-guides across the two L2s, and Steiger’s (1980) Z-tests to compare dependent correlations.

Results

Internal Consistency Reliability

Table 2 shows the Cronbach’s alpha reliability coefficients for the various multi-item scales of the present study. Most of the reliability coefficients are above (or very close to) the recommended .70 threshold; the only exception is the Visual style scale—the unexpectedly low Cronbach’s alpha coefficient of this scale warrants further exploration.

The Visual style variable was originally measured by five items, all adapted from two well-established questionnaires, the LSS and PLSPQ. As can be seen in the following list of the original five items we used, they were either taken directly from those surveys or were slightly modified to better suit the particular context:
Table 3 Correlations between the future self-guides and various criterion measures

<table>
<thead>
<tr>
<th></th>
<th>Ideal English Self</th>
<th>Ideal Mandarin Self</th>
<th>Ought-to English Self</th>
<th>Ought-to Mandarin Self</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intended effort in English</td>
<td>.68***</td>
<td>.47***</td>
<td>.24**</td>
<td>.04</td>
</tr>
<tr>
<td>English grades</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended effort in Mandarin</td>
<td></td>
<td>.67***</td>
<td></td>
<td>.59***</td>
</tr>
<tr>
<td>Mandarin grades</td>
<td>.42***</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p < .01; ***p < .001.

1. I understand better by reading instructions than by listening to instructions. (Adapted from PLSPQ: “I understand better when I read instructions.”)
2. I remember something better if I write it down. (Taken over from LSS.)
3. I use highlighters to help me as I learn. (Adapted from LSS: “I use color-coding to help me as I learn or work.”)
4. I learn better by reading what the teacher writes on the chalkboard. (Taken over from PLSPQ.)
5. I learn more by reading textbooks than by listening to the teacher. (Adapted from PLSPQ: “I learn more by reading textbooks than by listening to lectures.”)

Because the straightforward content and the tried-and-tested nature of these items served as an assurance of their validity, our main concern was the low reliability of the composite scale, which would reduce statistical results due to its lower sensitivity. Because of the significance of this scale in our study, we decided not to exclude it from the analyses, and as we shall see later in this section, it did produce important significant results. Interestingly, the two items that most reduced the scale’s reliability—and were thus excluded—were the first and the fifth ones, which specifically contrasted reading and listening skills.

Correlations Between the Future Self-Guides and Various Criterion Measures (H1)
The results identified strong, positive correlations between both sets of ideal self-guides (English and Mandarin) and learners’ effort in L2 learning (Table 3). This positive association was also borne out by the correlations with actual grades obtained, which, although understandably lower than the relationship with the self-report measures (as actual grades are also affected
Connections Between the Future Self-Guides and the Visualization-Related Measures (H2a)

Table 4 presents correlations and multiple correlations between the learners’ future L2 self images and the visualization-related measures (visual learning style and imagery capacity). The results confirmed that future self-guides are strongly associated with visual imagery: All the correlations in Table 4 are significant, and significant links were not only identified with the ideal L2 selves but also with the ought-to L2 selves in both target languages. Upon examining the joint effect of the two visualization-related variables by means of multiple correlations, we obtained coefficients that explain 14–21% of the shared variance, attesting to the fact that mental visualization is a notable component in learners’ imagined future L2 self identities.

Adding Auditory Learning Style Preference to the Paradigm (H2b)

Table 5 presents the correlations between auditory style and the future self-guides for both target languages. As the results indicate, similar to visual style, auditory style has significant positive associations with the self dimensions. We also computed multiple correlations in which we pooled the three sensory variables. The figures show that, by adding auditory style, the multiple
Table 5  Adding auditory style to the analyses: Correlations and multiple correlations

<table>
<thead>
<tr>
<th></th>
<th>Ideal English Self</th>
<th>Ought-to English Self</th>
<th>Ideal Mandarin Self</th>
<th>Ought-to Mandarin Self</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory style ((r))</td>
<td>.43***</td>
<td>.21**</td>
<td>.29***</td>
<td>.17*</td>
</tr>
<tr>
<td>Multiple correlation ((R))</td>
<td>.52***</td>
<td>.41***</td>
<td>.48***</td>
<td>.37***</td>
</tr>
</tbody>
</table>

(Visual style, Auditory style, and Imagery capacity)

\* \(p < .05\); ** \(p < .01\); *** \(p < .001\).

correlations further increased relative to the multiple correlations presented in Table 4, explaining as much as 27% of the ideal English self image.

The Language-Independence of Imagery Skills (H3)
In order to examine whether the learners’ skills in generating imagery affect distinct language selves differently, we computed Steiger’s (1980) \(Z\)-test results to test whether the correlations of imagery with the ideal and the ought-to self-guides as well as with the criterion measures are significantly different from each other across the two target languages. A consistent picture emerged: None of the \(Z\) values related to the correlations with overall imagery (i.e., combined visual style, auditory style and imagery capacity) were significant (ideal L2 selves: \(r_{Eng} = .52, r_{Mand} = .48, Z = .54\); ought-to L2 selves: \(r_{Eng} = .41, r_{Mand} = .37, Z = .64\); intended effort: \(r_{Eng} = .46, r_{Mand} = .37, Z = 1.35\); grades: \(r_{Eng} = .15, r_{Mand} = .25, Z = 1.23\)), indicating that these correlations did not show any language-specific difference. That is, imagery exerted its impact regardless of what the specific target language was.

Examining the Distinctive Nature of L2-Specific Self-Guides
Our final research question concerned the distinctiveness of self-guides related to different target languages. We present three sets of findings to explore this question. First, we carried out independent-samples \(t\)-tests to compare the corresponding self-guides, that is, the ideal English and Mandarin selves and the ought-to English and Mandarin selves. The results were significant for ideal English self \((M = 3.62, SD = .74)\) and ideal Mandarin self \((M = 3.38, SD = .81)\), \(t(171) = 3.52, p < .01\), effect size (eta squared) = .07, as well as for ought-to English self \((M = 3.67, SD = .79)\) and ought-to Mandarin self \((M = 3.01, SD = .74)\), \(t(171) = 12.34, p < .001\), effect size (eta squared) = .47.
Dörnyei and Chan  
Motivation and Vision

**Table 6** The contrasting motivational capacity of different language selves: Correlation coefficients and Steiger’s Z-scores (in parentheses) comparing the correlation pairs

<table>
<thead>
<tr>
<th></th>
<th>Ideal English Self</th>
<th>Ought-to English Self</th>
<th>Ideal Mandarin Self</th>
<th>Ought-to Mandarin Self</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English/Mandarin</td>
<td>.68/.23 (6.43***);</td>
<td>.47/.23 (3.12***);</td>
<td>.22/.67 (6.37***);</td>
<td>.21/.59 (5.13***);</td>
</tr>
<tr>
<td>(Z)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades:</td>
<td>.24/.17 (.86)</td>
<td>.04/.00 (.47)</td>
<td>.14/.42 (3.54***);</td>
<td>.05/.15 (1.20)</td>
</tr>
<tr>
<td>English/Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Z)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; ***p < .001.

These significant differences with moderate to large effect sizes indicate distinct language-specific magnitudes.

Second, the pairs of correlation coefficients presented in Table 6 show that the motivational capacity of the different self-guides is primarily exerted on the corresponding criterion measures. Where there is a significant difference between the magnitude of the correlations, the English self-guides impact English criterion measures more than Mandarin ones and, similarly, Mandarin self-guides impact Mandarin criterion measures more than English ones. (Because, as we have seen earlier, the correlations with actual grades were lower, their differences reached significance only for the Ideal Mandarin self.) These results point to distinct language-specific motivational functions.

Finally, we conducted two exploratory factor analyses, submitting to them all the items constituting the two ideal self scales (i.e., the scales for English and Mandarin) and the two ought-to self scales, respectively. Any distinctness of the L2-specific self images would be indicated by the English-related and the Mandarin-related items forming separate factors. Regarding the ideal self, a two-factor solution produced a clear and meaningful factor matrix (Table 7), with the items separating neatly into two clusters according to the two target languages. The ought-to self dimension presented a more complex picture (Table 8), because here a four-factor solution emerged, revealing that different aspects of the ought-to self dimension behaved differently in our sample regarding their relation to the two target languages. While the immediate *social expectations* coming from authority figures—which constitute the core of the classic ought-to self construct—did display distinct English and
Table 7  Exploratory factor analysis of the items constituting the ideal self scales (maximum likelihood extraction, oblimin rotation, pattern matrix; two-factor solution; factor loadings below .35 omitted)

<table>
<thead>
<tr>
<th>Ideal Self Factors</th>
<th>English</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can imagine myself participating in a debate in these languages. (English)</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>I can imagine myself being a very competent speaker of these languages. (English)</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>I can imagine myself writing e-mails in these languages fluently. (English)</td>
<td>.70</td>
<td></td>
</tr>
<tr>
<td>When I think of the future, I can imagine myself using these languages in a variety of ways. (English)</td>
<td>.43</td>
<td></td>
</tr>
<tr>
<td>I often imagine myself speaking these languages as if I were a native speaker of these languages. (English)</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td>I often imagine myself speaking these languages as if I were a native speaker of these languages. (Mandarin)</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>I can imagine myself being a very competent speaker of these languages. (Mandarin)</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>When I think of the future, I can imagine myself using these languages in a variety of ways. (Mandarin)</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>I can imagine myself writing e-mails in these languages fluently. (Mandarin)</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td>I can imagine myself participating in a debate in these languages. (Mandarin)</td>
<td>.54</td>
<td></td>
</tr>
</tbody>
</table>

Mandarin facets (Factors 1 and 4), such a language-based separation did not emerge with regard to two further aspects of the overall ought-to self construct: friends’ opinions (Factor 2) and the general sociocultural climate (i.e., the extent to which L2 proficiency is part of an educated social image within the social environment) (Factor 3).

Discussion

The results of this study produced interesting insights, providing unambiguous responses regarding all the issues raised about the ideal L2 self and partial answers regarding the ought-to L2 self. H1 stated that both the ideal and the ought-to L2 selves would be positively associated with the two criterion measures: self-reported intended effort and course grades, both for English
Table 8 Exploratory factor analysis of the items constituting the ought-to self scales (maximum likelihood extraction, oblimin rotation, pattern matrix; four-factor solution; factor loadings below .30 omitted)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I consider learning these languages important because the people I respect think that I should do it. <em>(Mandarin)</em></td>
<td>1.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studying these languages is important to me in order to gain the approval of my family. <em>(Mandarin)</em></td>
<td></td>
<td>.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have to study these languages, because, otherwise, I think my parents will be disappointed with me. <em>(Mandarin)</em></td>
<td></td>
<td></td>
<td>.32</td>
<td></td>
</tr>
<tr>
<td>I study these languages because close friends of mine think they are important. <em>(English)</em></td>
<td></td>
<td></td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>I study these languages because close friends of mine think they are important. <em>(Mandarin)</em></td>
<td></td>
<td></td>
<td>.70</td>
<td></td>
</tr>
<tr>
<td>People around me believe that I must study these languages to be an educated person. <em>(Mandarin)</em></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>People around me believe that I must study these languages to be an educated person. <em>(English)</em></td>
<td></td>
<td></td>
<td>.53</td>
<td>.43</td>
</tr>
<tr>
<td>Studying these languages is important to me in order to gain the approval of my family. <em>(English)</em></td>
<td></td>
<td></td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>I have to study these languages, because, otherwise, I think my parents will be disappointed with me. <em>(English)</em></td>
<td></td>
<td></td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>I consider learning these languages important because the people I respect think that I should do it. <em>(English)</em></td>
<td></td>
<td></td>
<td>.48</td>
<td>.52</td>
</tr>
</tbody>
</table>
and Mandarin. This was indeed confirmed for the ideal self images, thereby verifying the motivational power of the ideal L2 self. It is noteworthy that, unlike most past studies that did not use actual course achievement data as criterion measures, here the strength of the desired self images was also verified through their association with actual course grades.

On the other hand, the ought-to self variables presented mixed results: Similar to previous studies, they were found to correlate positively with intended effort for both target languages, but their associations with the course grades were nonsignificant. That is, while the participants perceived the external pressures on them as being valid and did intend to adjust their behavior accordingly, this intended effort was not manifested in their actual grades. The limited motivational capacity of the ought-to self that we have found here has been, in fact, also displayed by the majority of past studies in which the ought-to L2 self typically had weaker links with the criterion measures than the ideal L2 self (Csizér & Kormos, 2009; Csizér & Lukács, 2010; Taguchi et al., 2009). There is, thus, a tentative conclusion emerging from the existing body of research that, while externally sourced self-images (i.e., the images that are usually categorized under the rubric of the ought-to self) do play a role in shaping the learners’ motivational mindset, in many language contexts they lack the energizing force to make a difference in actual motivated learner behaviors by themselves.

H2 assumed that there will be positive correlations between the future self-guides and both the visualization-related and the auditory style variables. This hypothesis was fully confirmed in spite of the limited reliability—and thus sensitivity—of the visual style scale as discussed earlier. Future self-guides were found to be strongly associated with visual imagery, and the addition of auditory style to the paradigm further strengthened the sensory link with both the ideal and the ought-to L2 selves in both target languages. The investigation of this question was the main driver behind the current study, and the results in this respect are consistent and unambiguous. Furthermore, the salience of these phenomena is attested to by the fact that the emerging correlation coefficients are higher than the figures usually reported as meaningful in L2 motivation studies. We can therefore conclude that mental imagery is indeed associated with future self-guides, which justifies the use of the term “vision” when referring to them. This imagery is of a multisensory nature: Our data suggest that, rather than understanding imagery merely as a form of visualization, it should be considered a simulated mental experience that also involves the other senses, particularly auditory perceptions. Thus, a fitting way of defining vision is to view it as the sensory experience of a future goal state or, in other words,
a personalized goal that the learner has made his/her own by adding to it the imagined reality of the goal experience.

H3 assumed that imagery is such a fundamental aspect of desired future language selves that it is L2 independent, that is, it is not affected by the nature of the target language involved. This hypothesis was confirmed, indicating that when learners pre-live future language experiences (i.e., have language-specific visions), the mental imagery component constitutes an inherent part of the generic mechanisms of mental simulation and is not directly related to the specific content of the vision (i.e., the specific target languages in our case).

Finally, the question as to whether learners have several distinct desired possible self images of themselves or one broad vision for the ideal and one for the ought-to selves, respectively, with various facets, has produced a mixed response. Taking together the results of the three analyses conducted to address this issue, the findings are uniform in that the ideal L2 images associated with different target languages are distinct, which confirms in a more conclusive manner the indications of previous research findings. For example, in a study using cluster analysis to identify learner profiles, Csizér and Dörnyei (2005; see also Dörnyei, Csizér, & Németh, 2006) have found that coexisting ideal language images in a person may cause interferences with each other in that a positive disposition toward one language can go at the expense of another. In other words, different visions can compete for their place in the working self-concept. Further information supporting this claim has recently been provided in two studies by Henry (2010, 2011), whose results also suggest the presence of independent L2 selves. Among Swedish learners whose L2 was English but who were also studying an L3 or a fourth language (L4), and so on, the English-specific self-guides formed a dominant presence in the working self-concept, functioning as a reference or standard in evaluating the L3 or L4 selves. Thus, our results add to the growing consensus in the field of L2 motivation that coexisting ideal L2 self images constitute fairly distinct L2-specific visions, which can then interfere with each other both in a positive way (e.g., transferable linguistic confidence from one language experience to the other) or in a negative, demotivating manner (e.g., competition for space in the working self-concept).

With regard to ought-to self images, the picture is less clear. Our study indicates that when the two target languages that are studied simultaneously both receive general social support (which is the case for English and Mandarin in Hong Kong), some of the learners’ imported visions—in our case particularly those that came from friends or from the more general social
environment—may not divide clearly into language-specific clusters. In other words, while some language impact was clearly detectable in the expectations coming from authority figures, this was confounded by another categorization principle, the division of the external images according to their source. This makes intuitive sense because foreign language proficiency is often referred to in a collective manner by external sources (e.g., by the media), which weakens and can even diminish the boundaries between L2-specific visions. It would be enlightening in this respect to study situations where, unlike in Hong Kong, the institutional and popular support for two target languages is not balanced, like for example the situation in former Communist countries such as Hungary, where the compulsory study of Russian used to receive full state support without popular endorsement, in contrast to the study of English or German, where the pattern was reversed.

**Limitations of the Current Study**

Besides the predictable weaknesses of a study such as ours—that is, the limited generalizability of the results to the population from a specific sample that is not fully representative, the cross-sectional nature of a survey that takes only a snapshot in time, and the various weaknesses of self-report questionnaires as research instruments—a specific limitation of our investigation was that our measurement of the sensory variables was not ideal for two reasons. First, the visual style scale had limited reliability, which restricted the scale’s sensitivity and therefore likely depressed the correlation coefficients and potentially affected the multiple correlational results that included this variable. Second, the short, one-dimensional imagery capacity scale could not do full justice to the various dimensions and modalities subsumed by the concept of imagery ability. In addition, in light of the ambiguities that have surfaced with regard to the ought-to self, it would have been better to apply more elaborate scales targeting different types of external pressures separately—as has been done, for example, in the instrument used by Taguchi et al. (2009)—instead of using a single ought-to self scale. Finally, recent theorizing about motivation within a dynamic systems framework (e.g., Dörnyei, in press) has highlighted the limitations of perceiving the impact of learner characteristics in a linear manner, as is typically done when using inferential statistical procedures for data analysis. Yet, the highly consistent results emerging in our study with regard to the ideal language self do point to the existence of a broad and potent attractor state, which would offset the above limitations to some extent.
Conclusion

The results obtained in this study are consistent with past investigations of L2 motivation in identifying significant positive associations between desired language self-guides (particularly the ideal L2 self) and the learners’ L2-related learning effort and achievement. With regard to our study’s specific focus on mental imagery, future self-guides were found to be associated with salient imagery/visualization components, which justifies the use of the term “vision” when referring to them. It was shown that this vision is multisensory in nature, involving all the senses and not just visualization. An important characteristic of the imagery skills involved was their language-independent nature, pointing to the conclusion that L2-related mental imagery is part of the more generic mechanisms underlying human vision rather than a function of specific languages. On the other hand, our findings confirmed that different languages are associated with distinct ideal language selves, thus forming distinct L2-specific visions. Further research is needed to decide the extent to which this also applies to the less internalized ought-to selves.

With respect to the practical implications of this study, the main point to highlight here is that the learners’ sensory/imagery capacity has considerable pedagogical relevance because it is an important internal resource that can be intentionally harnessed (Sheikh, Skeikh, & Moleski, 2002; Taylor, Pham, Rivkin, & Armor, 1998). Dörnyei and Kubanyiova’s (in press) recent book-length overview of the topic provides ample evidence that imagery skills are trainable; therefore, imagery training and guided imagery have a realistic potential to enhance L2 motivation by helping students to generate personal visions supported by vivid and lively images and then to sustain this vision during the often challenging everyday reality of the language learning process. As mentioned earlier, there have already been promising attempts to develop visionary training programs (Fukada et al., 2011; Magid & Chan, 2012; Sampson, 2012) and teachers can also consult two available practical resource books for vision-enhancing classroom activities (Hadfield & Dörnyei, 2013; see also Arnold, Puchta, & Rinvolucri, 2007).

Final revised version accepted 25 December 2012

References


Appendix

Items measuring imagery capacity in the study:

- If I wish, I can imagine some things so vividly that they hold my attention as a good movie or story does.
- Sometimes images come to me without the slightest effort.
- When I am thinking, I often have visual images rather than thoughts in my mind.
- My daydreams are sometimes so vivid I feel as though I actually experience the scene.
- When reading fiction I usually have a vivid mental picture of the scene that has been described.