ORIGINAL ARTICLE

Bernhard Hommel · Beatrix Eglau

Control of stimulus-response translation in dual-task performance

Abstract In contradiction to stimulus-response- (S-R-) translation bottleneck models of dual-task control, stimulus processing in a primary task is affected by its compatibility with the response in a secondary, later performed task (Hommel, 1998a)– an indication of parallel S-R translation. Here we show that this backward-compatibility effect is independent of working-memory load, whether this is induced by an extra memory task (Experiment 1) or by increasing the number of S-R alternatives in the primary task (Experiment 2). However, backward effects occur even when the secondary task is no longer carried out (Experiment 3) and they are strongly affected by the inconsistency of previously used S-R mappings (Experiment 4). These findings suggest that S-R translation is (or can be) capacity-independent and automatic even under multiple-task conditions, and that it is mediated by direct S-R associations that emerge after only little practice.

The search for capacity limitations in human information processing has long been a central issue in cognitive psychology, for both theoretical and practical reasons. The assumption of two processing bottlenecks has received rather general agreement; one processing bottleneck associated with the selection of stimuli for further processing (selective attention) and the other associated with the translation of stimulus information into an appropriate response. Most evidence for this latter, response-related bottleneck, which our investigation aims at, comes from dual-task studies. In these studies, people are commonly asked to perform a primary response (R1) to a primary stimulus (S1) and secondary response (R2) to a secondary stimulus (S2). The stimulus-onset asynchrony (SOA) between S1 and S2 often varies, so that primary and secondary tasks overlap temporally to varying degrees. With very few exceptions, overlap strongly impairs performance on the secondary task, and the greater the overlap, the more so (i.e., the shorter the SOA; for overviews see Meyer & Kieras, 1997; Pashler, 1994). This has been taken to indicate a structural bottleneck in information processing (though not without exception: Meyer & Kieras, 1997). Peripheral stages do not seem to be responsible, as dual-task costs are observed under conditions where sensory or motor cross talk can be excluded (e.g., De Jong, 1993; Pashler & Johnston, 1989). However, dual-task costs do interact with S-R compatibility (McCann & Johnston, 1992), this pointing to the S-R translation stage. Indeed, several authors have suggested that only one stimulus can be translated into a response at a time, so that secondary-task translation cannot begin before primary-task translation is completed (De Jong, 1993; Fagot & Pashler, 1992; McCann & Johnston, 1992; Welford, 1952).

Although most researchers agree with the idea that selecting a response provides some kind of bottleneck (the debated questions being, rather, whether this is the only bottleneck and whether it is structural or functional), its specific characteristics and possible functions are not well understood. Part of the problem has to do with the lack of conceptual clarity with respect to the terms *S-R translation* and *response selection*. Earlier approaches, like that of Smith (1967) or Welford (1952), assumed that responses are selected *by* translating a stimulus into a response (cf., Hommel, 2000a). Accordingly, the two terms were, and still are, treated as equivalent (Pashler, 1998, p. 277). Logically, however, they are not.

The term "translation" refers to the act of transforming some code into some other code of a different "language" or reference system (obviously reflecting the widespread belief of a fundamental incommensurability of perceptual and action-related codes; Prinz, 1984, 1990) or, in neural-network terms, to activating one code as a function of the activation level of another.

B. Hommel (✉)
Leiden University, Cognitive Psychology Unit,
Postbus 9555, 2300 RB Leiden, The Netherlands
E-mail: hommel@fsw.leidenuniv.nl

B. Hommel · B. Eglau
Max Planck Institute for Psychological Research,
Munich, Germany

Assuming a dual-task bottleneck associated with this process would imply that S2-R2 translation could not start while S1-R1 translation is going on (see Figure 1A). Indeed, such a model has been proposed by Welford (1952),[1] who claimed that S2 needs to be stored until R1 selection (and, possibly, even the processing of R1-related feedback) is completed, and only then used to identify R2. More modern versions of this type of model are discussed in Fagot and Pashler (1992) and McCann and Johnston (1992).

If one considers S-R translation as a discrete, single-channel process, there is no need to distinguish it from response selection: the proper stimulus directly specifies the appropriate response, which is then carried out. Yet, from a more continuous, parallel view, as propagated by Eriksen and Schultz (1979) and others, translation and selection are rather different functions: the former serves for activating response codes by applying context-specific translation rules and/or by using overlearned S-R associations, whereas the latter makes sure that a choice is made among the possibly many competing response codes (Berlyne, 1957). Assuming a dual-task bottleneck associated with response selection proper would not necessarily place restrictions on the parallelism of S-R translation. That is, S1-R1 translation and S2-R2 translation may very well overlap but maybe only one response decision can be made at a time (see Figure 1). Recent versions of this kind of model have been considered by Hommel (1998a), Jolicœur, Tombu, Oriet, and Stevanovski (2002), Logan and Schulkind (2000), Pashler (1993), and Ruthruff, Johnston, and Van Selst (2001).

Preliminary evidence in favor of a selection model may be taken from observations in single tasks involving stimuli that signal more than one response at a time. The most prominent examples are the Stroop task – where color words can appear in incongruent colors, the Simon task– where the critical stimulus feature can appear in a location corresponding to the incorrect response, or the flanker task – where target symbols can appear among flankers that signal another response (for an overview, see Hommel, 2000a). Inasmuch as these effects reflect response conflict, they demonstrate that more than one stimulus feature at a time can be translated into response activation, which again is inconsistent with a pure translation model (Fagot & Pashler, 1992). However, what these effects actually show is only that *one* given set of S-R translation rules can be concurrently applied to more than one piece of information. Yet, in standard dual-task situations it is not one but two sets of rules that need to be applied. Accordingly, there may still be a S-R translation bottleneck arising from strict limitations in the number of rules or rule sets that can be stored in working memory or that can be concurrently accessed (McCann & Johnston, 1992). Moreover, both the Stroop and the flanker effects have been claimed to in-

Fig. 1 A: Locations of possible S-R translation bottlenecks and response selection bottlenecks in the information flow. B: Two possible ways of S-R translation. According to the *transient-link model*, multiple S-R mapping rules can be implemented in working memory and used for parallel, automatic S-R translation (broken lines), perhaps in addition to serial, controlled S-R translation. According to the *permanent-link model*, automatic translation runs in parallel via acquired, direct S-R associations outside working memory (straight lines), perhaps in addition to serial, controlled S-R translation

clude a substantial component that goes back to conflict between stimulus codes rather than response conflict (e.g., Kornblum, Hasbroucq, & Osman, 1990), and the Simon effect with its tight association between spatial stimulus and response codes may not be the best example of parallel *rule* use anyway.

More direct evidence for parallel S-R translation has been reported only recently by Hommel (1998a), Stoet and Hommel (1999), and Logan and Schulkind (2000). Hommel (1998a) had people perform a manual primary task and a vocal secondary task, and varied the compatibility between the secondary response and either the primary response (R2-R1 compatibility) or the primary stimulus (R2-S1 compatibility). Performance in the primary task was strongly affected by either type of compatibility. For instance, responding to the color red was faster if the secondary response required saying "red" rather than "green", and vice versa. Interestingly, this kind of *backward compatibility* effect[2] was found even

---

[1]Curiously enough, the term "S-R translation" is commonly traced back to Welford's (1952) paper (e.g., Pashler, 1998, p. 277), where, however, it is not mentioned once.
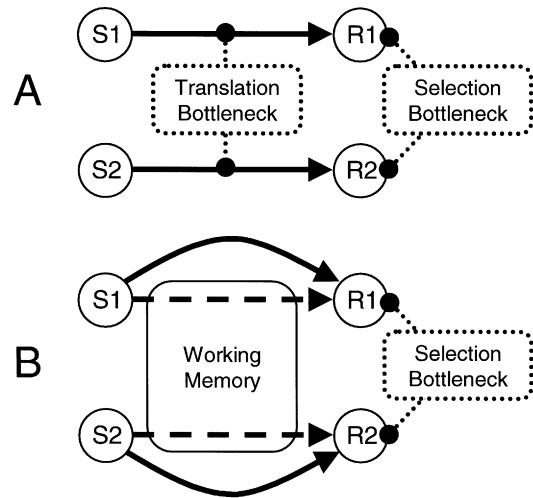
[2]Here and in the following we focus on the effect of the compatibility relationship between the secondary response and the primary stimulus on processing the primary stimulus, that is, on what we call backward compatibility. Of course, this relationship is likely to affect performance on the secondary task as well: When people react to the identity of colored letters by uttering color names, one would expect this reaction to be facilitated if the stimulus color corresponds to the correct response (e.g., red → "red") than if it does not (e.g., green → "red"), and the findings of Hommel (1998a) show that this is exactly what happens. However, the presence of such a nonspatial Simon effect (for an overview, see Lu & Proctor, 1995; Simon, 1990) is neither surprising nor of theoretical interest for present purposes. Therefore, we will present the relevant data and analyses to the interested reader but do not further discuss the respective outcomes.

though R2 was performed about half a second or more after R1, hence, there was a pronounced dual-task cost. However, the processing of S1 or R1 could only be affected by R2 if S2 was translated into R2 *before* S1 or R1 processing was completed, which strongly suggests that S2-R2 translation was *not* delayed by primary-task processing. Stoet and Hommel (1999) had people perform two responses that were signaled in reverse order, that is, S1 preceded S2 but R2 was performed before R1. In one condition (Experiment 3, short delay), preparing R1 before R2 was discouraged by presenting S1 only 100 ms before S2, and by leaving S1 on the display until R2 was initiated. Indeed, R1 was initiated 72 ms later than in conditions with very long S1 preview, indicating that R1 was not (strongly) prepared. Nevertheless, R2 was performed faster if it was compatible with R1, which shows that R1 must have been automatically activated by S1 to some degree. Logan and Schulkind (2000) used a similar design as Hommel (1998a) with a primary task requiring, for instance, a letter-digit classification of S1 and a secondary task requiring a letter-digit classification of S2. Again, R2-R1 compatibility strongly affected primary-task performance, that is, R1 was faster if it was compatible with R2. Obviously, S2 was translated into R2 before R1 selection was completed, which strongly supports a parallel-translation model.

Thus far, the available evidence suggests that S-R translation under dual-task conditions works like a prepared reflex (Hommel, 2000a), that is, once implemented the S-R rules more or less automatically translate the given stimulus information into the corresponding response activation. The purpose of the present study was to investigate *how* automatic this translation really is and how it may work. According to the *transient-link model* considered by Hommel (1998a), instructed S-R rules are stored in some kind of transient working memory, where they can be accessed by any stimulus defined in the "S" part of the rule, which is then translated into response activation (cf., Fagot & Pashler, 1992; McCann & Johnston, 1992; Meyer & Kieras, 1997). The capacity of this memory is likely to have some upper limit, but it should hold at least four rules (the number used in many dual-task experiments) concurrently. As soon as a stimulus is presented, the corresponding response or, more precisely, the codes representing this response will be activated. In a dual-task situation (with short SOAs) this means that the codes of two responses will be active at the same time, so that some further action is needed to organize the response activation in ways that two separate action plans can be formed (cf., Stoet & Hommel, 1999). Important implications of this kind of model are that the efficiency of S-R translation should be affected by working memory load, as tested in Experiments 1 and 2, and that rules should be exchanged, replaced, and deleted very quickly – a prediction tested in Experiments 3 and 4.

According to the second alternative discussed by Hommel (1998a), the *direct*- or *permanent-link model*, repeatedly applying the same S-R mapping rules, may lead to the formation of direct, permanent associations between the codes representing stimulus and response features, or to the accumulation of multiple memory traces of the S-R episodes[3](Logan, 1988; Logan & Schulkind, 2000). In either case, stimulus presentation would lead to immediate response activation without any direct involvement of working memory or the mapping rules stored therein. So, whether or not there are capacity limitations as to the number of rules held in working memory or the number of processes accessing those rules at a time, response activation itself (or at least part of it) would not underlie those limitations, simply because it proceeds (at least in part) along routes that are independent from transient memory structures. Accordingly, this model predicts no dependency of S-R translation effects on memory load but measurable inertia of the S-R associations. That is, S-R associations should outlive the situation they were needed for.

It should be emphasized that transient- and permanent-link models need not be mutually exclusive theoretical alternatives. Models of skill acquisition commonly distinguish learning phases that differ in their degree of automaticity (e.g., Anderson, 1982; Fitts, 1964), and it may well be that S-R links are transient at the time they are first implemented but become permanent through practice. It may even be that transient and permanent links co-exist, as assumed by dual-process models (e.g., Ach, 1910; Logan, 1988). Accordingly, we do not consider the transient-link and the permanent-link model to be comprehensive accounts of all kinds of S-R processing taking place under dual-task conditions; we only attempt to characterize those cognitive processes that mediate backward-compatibility effects. We will come back to these issues in Experiments 3 and 4, and in the General Discussion.

Experiment 1

The transient-link model is based on the idea that automatic S-R translation proceeds along temporary S-R associations implemented in a transient storage medium, in working memory. This means that backward-compatibility effects should be sensitive to manipulations that increase the load of, or competition within working memory: Increasing load should impair the degree to which S-R mapping rules are prepared, which should decrease the flow of information from S to R codes. The

---

[3]Traditional S-R learning models hold that S-R learning leads to the formation of direct, increasingly strengthened connections between stimulus and response codes (e.g., Thorndike, 1927), whereas Logan (1988) proposes an accumulation of multiple traces of S-R episodes. These two approaches explain the effect of practice in different ways, by referring to activation that spreads from stimulus to responses codes versus a stimulus-induced race between stored instances. However, for present purposes these differences do not matter and our findings will not discriminate between the two views. Therefore, when we speak of S-R associations, this should not be taken to imply any justifiable preference for one or the other approach.

less flow there is, especially from S2 to R2, the less impact R2 should have on S1 processing, hence, the smaller the backward-compatibility effect. In contrast, the permanent-link model assumes links outside working memory and therefore predicts no such relationship, hence, independence of backward-compatibility effects from working-memory load.

Logan (1978, 1979, 1980) introduced a method to load working memory in ways that are likely to affect S-R translation processes. Before performing a trial, participants were presented a list of digits, which they were to reproduce after the trial in correct order. The number of digits in a list was varied to introduce low and high degrees of memory load. Although memory load typically has a strong effect on performance in both the primary task and the memory task itself, it does not interact with a whole number of manipulations of perceptual factors (e.g., stimulus brightness or quality), attentional factors (e.g., array size or cuing), and decisional or response-related factors (e.g., yes-no decision or S-R compatibility) (for an overview, see Logan, 1980). Interestingly, however, memory load does interact with some factors that are likely related to S-R translation, such as response-set size (Logan, 1979), S-R mapping-rule complexity (Logan, 1980), or stimulus predictability (Keele & Boies, 1973). This suggests that, as Logan (1980) has pointed out, the digit-memory task uses up capacity from the same system or medium that prepares and holds S-R mapping rules, possibly by interfering with rehearsing the rules before the trial (Pashler, 1994).

The digit-memory task provides a means to test whether backward-compatibility effects depend on working-memory capacity. We therefore combined Hommel's (1998a) dual-task design, which we know produces the relevant backward effects, with Logan's memory task, this way creating a rather demanding triple task. In particular, we compared the effects of a rather trivial version of the digit task (remembering the order of two known digits) and a very difficult version (remembering the order of eight known digits) on the size of backward-compatibility effects.

Method

Participants

Twenty-four adults were paid to participate in a single session of about 90 min. They reported having normal or corrected-to-normal vision and were not familiar with the purpose of the experiment.

Apparatus and Stimuli

The experiment was controlled by a Hewlett Packard PC, interfaced with an A/D card for sampling of vocal input, attached to one monitor for stimulus presentation and another for online monitoring of vocal errors by the experimenter (for further details, see Hommel, 1998a). Manual responses were performed by touching the left or right of two touch-sensitive metal plates with the index finger of the left or right hand, respectively. Participants wore headphones with an integrated microphone, which registered the vocal responses "rot" and "grün", the German words for "red" and "green". Responses for the memory task were typed by the participants into the number block of the computer keyboard. The stimuli appeared at the center of the black screen. A white asterisk served as fixation mark. Stimuli for the primary and the secondary task were the uppercase letters *H* and *S* (S2), presented in red or green color (S1). For the memory task, the digits 0 and 9 were used as the 2-item set and the digits 1–8 as the 8-item set. Digits were shown in white, one at a time. From the viewing distance of about 60 cm, each symbol measured about .30 x .40.

Design

Each session consisted of a practice block for the memory task, a practice block for the manual-vocal dual task, and an experimental test block with all three tasks combined. The memory practice block included 10 trials, five for each item set, presented in alternation. The dual-task practice block consisted of 40 trials, composed of 10 replications of each of the four combinations of letter identity and color, randomly intermixed. In the test block, participants worked through four 40-trial miniblocks, each being composed of five replications of each of the eight combinations of letter identity, color, and memory load (2-item set vs. 8-item set), all randomly intermixed. That is, there was a total of 160 experimental trials, 20 for each condition. Primary, manual responses were assigned to stimulus color; secondary, vocal responses to letter identity. This yields four possible mapping combinations that were counterbalanced across participants.

Procedure

A memory-practice trial began with the presentation of an alerting message ("the next sequence of digits is:") for 1 s. Next, the randomly ordered items of the 2-item or the 8-item set were displayed, one by one, for 750 ms each. After deletion of the last item and a 7-s blank interval, a prompt requested the entry of the two or eight digits in correct order. There was no time limit and corrections were possible until the entry was confirmed by pressing the Enter key. One second later feedback about the number of correct entries (i.e., digits in correct positions) was presented for 2 s, and then the next trial started.

In the dual-task practice block each trial started, after an intertrial interval of 1 s, with presentation of the fixation mark for 1 s. After a blank interval of 250 ms the stimulus, a red or green *H* or *S*, was presented until both responses were made or 2.5 s had passed. Participants made a manual left-right response (R1) to stimulus color and a vocal response ("red" or "green"; R2) to the letter, according to the instructed mapping rule. They were required to perform the two responses in strict serial order and an error message was displayed (and the corresponding responses were counted as incorrect) if the primary response did not lead over the secondary by at least 50 ms. In case of an order error, a premature (RT < 150 ms) or omitted response (RT > 2 s), an incorrect keypress, or an incorrect vocal response (as determined by the experimenter during the intertrial interval), a brief error message was displayed on the screen, accompanied by a beep. The corresponding trial was recorded and repeated at some random position in the remainder of the block.

Experimental trials were as in the dual-task practice block, except that the manual-vocal task was now combined with the already practiced memory task. That is, before the fixation mark was presented, the randomly ordered memory set was displayed, just as in the memory practice block, and participants were required to hold this set (i.e., the items in their correct order) in memory. After the secondary response, and only if both responses were correct, participants were requested to type in the memorized items in the correct order, as in the memory practice block. Again,

entry was unspeeded and corrections were possible, but no feed-back was given.

Results

Premature responses accounted for 0.9% of the trials, response omissions for 4.4%, and order errors (secondary before primary response) for 2.2%. These trials and trials where one correct response was accompanied by a choice error in the other (concomitant errors: 1.2% of the primary and 2.3% of the secondary responses) were excluded from analyses in this and all other experiments of this study. From the remaining data, mean RTs and percentages of (choice) errors (PEs) were computed for each response as a function of compatibility (primary response to red stimulus/secondary response "red", primary response to green stimulus/secondary response "green") or incompatibility (primary response to red stimulus/secondary response "green", primary response to green stimulus/secondary response "red") between "primary" stimulus and secondary response (see Table 1). An alpha level of 5% was used for all statistical tests.

In the memory task, performance was better with low memory load (1.9 of 2 items, or 96% correct) than with high memory load (5.6 of 8 items, or 70% correct), $F(1,23) = 103.64$. The effect of compatibility and the interaction approached significance ($p < .1$), reflecting that memory performance was slightly impaired in incompatible as compared to compatible trials in the low-load condition (95% vs. 97%), but did not vary with compatibility in the high-load condition (70% vs. 70%).

In the omnibus ANOVA on RTs with response, memory load, and compatibility as within-subjects factors, all three main effects were significant: response, $F(1,23) = 162.06$, memory load, $F(1,23) = 654.96$, and compatibility, $F(1,23) = 31.91$, indicating longer latencies with secondary versus primary responses, with high versus low memory load, and with an incompatible versus compatible relationship between primary stimulus and secondary response. Also significant was the interaction of response and compatibility, $F(1,23) = 8.88$, due to larger compatibility effects on secondary versus primary responses. Although the interaction of memory load and compatibility approached the

significance criterion ($p = .127$), this indicated an *increase* rather than a decrease of the compatibility effect under high memory load. Separate ANOVAs on primary and secondary responses yielded the same effects as found in the omnibus analyses (except, of course, the effect of response).

PEs showed a very similar pattern as RTs, with significant main effects of memory load, $F(1,23) = 9.58$, and compatibility, $F(1,23) = 12.97$. These effects were also reliable in separate analyses of the two tasks.

Discussion

First of all, pronounced backward-compatibility effects were observed, that is, primary-task performance was better if S1 was compatible than incompatible with R2. This means that, even though R2 followed R1 by about half a second on average, it must have been activated to a considerable degree before the processing of S1 was completed. If so, translating S2 into R2 cannot have been delayed until S1-R1 translation was completed, which in turn suggests automatic S-R translation. These findings are fully consistent with the previous observations of Hommel (1998a) and Logan and Schulkind (2000) in demonstrating that the backward-compatibility effect is replicable even under the more demanding triple-task conditions employed here.

More interesting in the present context, there was no indication that the backward effect would decrease with increasing memory load – if anything, there was an (albeit unreliable) *increase*. Although this is a statistical null effect, it is obtained vis-à-vis a very pronounced effect of memory load upon performance in all three tasks. Therefore, we think that our load manipulation should have been strong enough to reveal possible impairments of S2-R2 translation, so that the absence of signs of such impairments can be taken as preliminary evidence of the independence of S-R translation from working memory capacity. This observation is inconsistent with predictions from a transient-link model as sketched above and, rather, points to a permanent-link account.

As an aside, it is interesting to contrast our finding of additive effects of working-memory load and backward compatibility, a measure of S-R translation efficiency,

Table 1 Mean Reaction Time (RT, in ms) and Percentage of Errors (PE) for Manual (Primary) and Vocal (Secondary) responses in Experiment 1 according to memory load (2 vs. 8 items in Experiment 1; 2 vs. 4 S-R alternatives in Primary Task of Experiment 2) and compatibility between Primary Stimulus and Secondary (Vocal). RT compatibility-effect sizes (Δ) in far right column

Experiment	Response	Memory Load	Compatible		Incompatible		
			RT	PE	RT	PE	Δ
1	Primary	low	795	0.9	940	4.7	145
		high	867	2.2	1047	5.4	180
	Secondary	low	1276	0.8	1467	2.8	191
		high	1354	1.4	1582	3.7	227
2	Primary	low	656	2.7	708	6.1	52
		high	853	4.4	906	5.8	53
	Secondary	low	1170	1.9	1251	4.7	81
		high	1351	1.7	1437	4.4	86

with those of Logan (1979), who varied memory load and response-set size, a factor that is also commonly associated with S-R translation. Logan's participants needed six days of practice before the load-by-set interaction gave way to additivity, this posing the question of why we were able to obtain additivity within a single session. The crucial difference in our view is likely related to the way the low-load condition was operationalized in the two studies. We had subjects retain 2 or 8 items, which leaves item number as the only difference between low- and high-load conditions. In Logan's study, however, memory sets of 0 and 8 items were compared, that is, the whole memory task was dropped under low-load conditions. This can be expected to have introduced more differences between conditions, such as the general preparation level or the degree to which cognitive capacities were distributed among the sub-tasks. If so, it is reasonable to assume that it was these non-specific factors that were responsible for the super-additivity in the earlier phases of Logan's study– factors that are unlikely to have played a role in our Experiment 1. In other words, the practice effects observed by Logan (1979) seem to reflect changes in the handling and scheduling of sub-tasks rather than in the way stimuli are translated into responses.

Experiment 2

Experiment 2 was conducted for two reasons. First, the evidence against the transient-link model gathered in Experiment 1 rests entirely on the absence of an interaction, hence, on a statistical null effect. Given the strong main effect of memory load and the tendency toward an interaction in the nonpredicted direction, this is not too disturbing, but we did wish to provide additional, converging evidence to bolster our conclusion.

Second, one may argue that the way we manipulated working-memory load in Experiment 1 was not sufficiently specific. There is broad agreement that working memory can be subdivided into relatively independent, content-specific subsystems, such as the articulatory loop and the visuo-spatial scratch-pad (Baddeley & Hitch, 1974) or separate phonological, visual, and spatial stores (Jonides & Smith, 1997), yet there is no consensus as to where in those stores S-R mapping rules are located. Accordingly, we cannot be sure that the capacity taken by our memory task really belonged to the part of working memory holding the S-R mapping rules. To circumvent these problems in Experiment 2 and to make sure that the loaded system is the one that does contain the S-R rules, we manipulated working-memory load by varying the number of S-R mapping rules required for the primary task. In particular, we contrasted the standard design with two primary-task and two secondary-task S-R alternatives (low-load condition, implying four S-R rules to be used) with a high-load condition comprising four primary-task and two secondary-task S-R alternatives (implying six S-R rules to be used).

Method

Twenty-four adults were paid to participate in a single session of about 45 min. They fulfilled the same criteria as in Experiment 1.

Apparatus and stimuli were as in Experiment 1 with the following exceptions. A new response board was constructed with four touch-sensitive metal plates to be touched with the middle and index fingers of the left or right hand. In the 2-alternative response-set section, responses were performed by touching the two inner plates with the index finger of the left or right hand, respectively. In the 4-alternative section all four plates were used. The stimuli for the secondary task were the uppercase letters H and S, as in Experiment 1. The colors used for the primary task depended on the number of manual response alternatives. In the two-alternative section, left- and right index-finger responses were signaled by the colors red and green. In the four-alternative section, the mapping for the two index fingers remained unchanged, and the additional middle-finger responses were signaled by the colors blue and yellow. As no memory task was presented, the digit stimuli were not used.

A session consisted of two sections, one with two and the other one with four manual response alternatives (and stimulus colors). Each section comprised a practice block composed of five replications of each of the four (in the 2-alternatives section) or eight (in the 4-alternatives section) combinations of letter identity and color. In the test block, participants worked through four 20-trial or 40-trial miniblocks (in the 2- and 4-alternatives section, respectively), each being composed of five replications of each of the four or eight combinations of letter identity and color, all randomly intermixed. That is, there was a total of 240 experimental trials, 80 in the 2-alternative section and 160 in the 4-alternative section, and 20 for each condition. Primary responses were again assigned to stimulus color, secondary responses to letter identity. The possible mapping combinations that were counterbalanced across participants according to a latin square design, except that the colors red and green were always mapped onto the two inner keys (i.e., the index-finger responses). Half of the participants began with the 2-alternative condition and the other half with the 4-alternative condition. The procedure was as in the dual-task practice block of Experiment 1.

Results

The data were treated as in Experiment 1. Data from middle-finger responses were excluded, because these were only available for the 4-alternative section. That is, the tabled means (see Table 1) and all analyses are based exclusively on the data from the two index-finger responses and, hence, on conditions with stimulus colors red and green. The omnibus ANOVA of RTs with response, number of stimulus-response alternatives, and compatibility as within-subjects factors yielded three main effects: response, $F(1,23) = 582.14$, alternatives, $F(1,23) = 53.21$, and compatibility, $F(1,23) = 41.44$, indicating slower responses with secondary versus primary responses, with four versus two alternatives, and with an incompatible versus compatible relationship between S1 and R2. As in Experiment 1, the interaction of response and compatibility was significant, $F(1,23) = 8.30$, due to larger compatibility effects on secondary versus primary responses. The interaction of alternatives and

compatibility was far from significance ($p > .9$), as was the three-way interaction with response ($p > .7$). Separate ANOVAs on primary- and secondary-task data yielded the same effects as the omnibus analyses. The PEs followed a similar pattern as the RTs, but only the main effects of response, $F(1,23) = 4.92$, and compatibility, $F(1,23) = 22.66$, were significant. The same effects were obtained in separate ANOVAs on data from the two tasks.

Discussion

The results are clear in showing that, first, the backward-compatibility effect was successfully replicated and, second, its size did not vary with the number of S-R rules in the primary task. This null effect is again obtained vis-à-vis a pronounced main effect of our load manipulation, which here was more than double the size of the already substantial load effect in Experiment 1. That is, even if one considers possible contributions of load-unrelated, pure stimulus and/or response factors (due to the increase of stimulus and response sets that accompanied the increase of the number of rules), there should have been plenty of variance for an impact on S2-R2 translation to account for. Yet, again no hint was obtained that backward-compatibility effects might depend on load.

Although such a result does not conform to predictions from a transient-link account, it is worthwhile to consider whether and how it might be saved. A possibility would be that the load manipulation affected S-R links in primary and secondary tasks to comparable degrees. Thus, it may well have been that high load did weaken S2-R2 links, as the transient-link approach predicts, but as they now competed with S1-R1 links that were weakened themselves, the resulting backward-compatibility effect remained unchanged in size. The apparent problem with this account is that it would predict load-induced costs in either task, which amounts to an over-additive interaction of response and load. Yet, not only was this interaction far from significance, $F(1,23) < .32$, the numerical load effect was even smaller in the secondary versus the primary task (see Table 1). That is, the load effect on the primary task merely propagated to the secondary task, which itself apparently did not contribute anything to it. But if it was not affected, there is no point in assuming that the S-R links it works with were weakened. This leaves two possible conclusions: Either subjects for some reason sacrificed precious working-memory capacity for optimizing the secondary task, thereby accepting pronounced, noticeable costs in primary-task performance; or the translation processes underlying backward-compatibility effects did not require such capacity, as the permanent-link approach suggests. In view of instructions stressing the greater importance of the primary task and of secondary-task RTs that suggest that the instructions worked, we consider the latter possibility more plausible than the former. Thus, inasmuch as backward-compatibility effects reflect S2-R2 translation, we take Experiment 2 to provide further evidence that the underlying translation processes do not draw on working-memory capacity. Given that very different load manipulations yielded comparable outcomes in Experiments 1 and 2, this seems to be true however working memory is defined– as long as some capacity limitation is assumed– and whatever the internal structure of this memory might be.

Comparing our first two experiments it is interesting to note that the compatibility effect shrank to about one-third of its size from Experiment 1 to 2. One explanation would be that even if the frequency of occurrence of individual S-R pairings is constant, forming a particular S-R association gets more difficult the more stimuli and responses are coded in the same context and, thus, the more other S-R associations are developing. If so, one might expect that the order of the 2- and 4-alternative sections played a role, such that beginning with the 2-alternative section would have allowed forming stronger associations than beginning with the 4-alternative section with its higher number of competing S-R relations. Indeed, when we entered section order as an additional factor in the analyses of the RTs, we found an interaction between order and compatibility that approached significance ($p = .07$). Participants beginning with the 2-alternative section had a compatibility effect of 62 ms that in the following 4-alternative section increased to 75 ms. Although these numbers are still smaller than those in Experiment 1, they do approach the level of compatibility effects observed under various conditions in the following experiments. However, participants beginning with the 4-alternative section had a compatibility effect of only 32 ms that in the following 2-alternative section increased to a mere 43 ms. This pattern suggests that automatic S-R translation (as indexed by compatibility effects) is not hampered by the *presence* of other associations (indeed, order did not interact with load, $p > .3$, or with load and compatibility, $p > .6$) but by the concurrent *formation* of such associations. However, once these associations are sufficiently strengthened, translation is independent of any working-memory load.

Taken together, Experiments 1 and 2 provide evidence favoring a permanent-link over a transient-link model of automatic S-R translation. Once the task-relevant S-R links are established, S2-R2 translation does not depend on, and is therefore presumably not mediated by, transient S-R rules temporarily stored in memory. Rather, it seems that relatively permanent S-R links – overlearned S-R associations or accumulated S-R episodes – are responsible for fast and effortless activation of secondary responses. Experiments 3 and 4 were conducted to test some implications of this view more directly.

Experiment 3

A crucial difference between the transient- and the permanent-link model concerns the role of S-R learning.

According to the transient-link model, no such learning is required to account for backward-compatibility effects, as these effects are mediated by temporarily stored S-R rules. According to the permanent-link model, however, backward effects are mediated by S-R associations that emerge through practice. An obvious test of this idea would see whether backward effects get stronger with practice, that is, whether they increase from early to later trials. Accordingly, Hommel (1998a) analyzed backward effects from a couple of experiments as a function of practice, but found no evidence of a systematic relationship between effect size and practice. However, *a priori* it is difficult to determine how much practice is needed to produce backward effects; in principle, a single trial per condition may be sufficient. If so, and given that measuring backward effects requires at least four trials per condition and a reasonable number of replications to get a reliable estimate, it would seem hopeless to look for fine-graded practice effects. Therefore, we used a somewhat different strategy in Experiment 3 by studying the after-effects of having acquired S-R links in a previous task.

If repeatedly performing a secondary task really leads to the formation of long-term links between the respective stimulus and response codes, as the permanent-link model holds, and if it is these links the backward effects are mediated by, those effects should be obtained even if performing the secondary task is no longer required. Accordingly, we gave subjects practice of half a session on our standard dual task and then asked them to stop performing the secondary task for the rest of the session. That is, participants were from this moment on no longer required to maintain or remember the S-R mapping rules of the secondary task. If under these conditions the compatibility between the (no longer performed) secondary response and the primary stimulus still mattered, this would provide further support for the idea of permanent S-R associations capable of mediating automatic S-R translation.

Yet, even if this could be demonstrated, the amount or efficiency of translation may nevertheless be moderated by the degree of preparation for processing the stimuli and or performing the responses involved in the particular S-R rules. Preparing for those stimuli and responses may prime their cognitive representations, which again may facilitate the flow of activation between these representations. To test this possibility, we investigated four groups of subjects that in the second, fade-out phase of the experiment differed with respect to their preparation for processing the stimuli and responses that previously belonged to the secondary task. In the *no preparation* (NP) group we merely introduced an infrequent nogo stimulus that signaled omitting the manual response. That is, in a few trials S2 was not an *H* or *S* (which now no longer required a response) but the letter *X*, which indicated that no response would be required in that trial. This manipulation was intended to make sure that subjects did not ignore the shape of S2 altogether without requiring any reaction to the former S2. In the *stimulus preparation* (SP) group nogo letters were more frequent and more diverse, and subjects were instructed to perform the manual response only in case of an *H* or *S*; i.e., now it was the go trials that were explicitly defined. In the *response preparation* (RP) group the stimulus was sometimes surrounded by a frame, in which case subjects should vocally name the color of the stimulus instead of pressing a key; a manipulation thought to prevent deactivation of vocal responses in the second phase. Finally, in the *stimulus and response preparation* (SRP) group the manipulations used in groups SP and RP groups were combined to foster preparation of both the stimulus and the response parts of the previous S2-R2 mapping rules.

The predictions with regard to the group manipulation depend on the degree of automaticity or context-independence of S-R links. If S-R translations were really completely automatic, the four preparation groups would be expected to produce comparable backward-compatibility effects; that is, the compatibility effect should not interact with group. However, if the efficiency of translation does depend on holding the stimulus or response part, or both, of the acquired S-R association in some state of preparation, compatibility effects should be more pronounced in groups SP, RP, and/or SRP, respectively. Hence, differences between the compatibility effects in the four groups should tell us what part(s) of the associations are necessary to prepare for S-R translation to take place.

Method

Eighty adults were paid to participate, 16 in group NP, 20 each in groups SP and RP, and 24 in group SRP. They fulfilled the same criteria as in Experiment 1. Apparatus and stimuli were as in Experiment 1, except that the stimuli from the memory task were not used and some additional stimuli were introduced to signal catch trials (see below). A session consisted of two sections.

The first dual-task section comprised five 20-trial blocks, that is, 100 experimental trials (there were no practice trials). Each block was composed of five replications of each of the four combinations of letter identity and color, randomly intermixed. Primary responses were assigned to stimulus color, secondary responses to letter identity, and the possible mapping combinations were counterbalanced across participants. The further procedure was as in the 2-alternative section of Experiment 2.

In the second section only the previously primary, but not the secondary, vocal task was to be performed, turning the dual-task into a single-task situation. The S1-R1 (color-key) mapping remained unchanged. Each of the five blocks in the second section included extra trials that differed between the four experimental groups. In group *NP* one catch trial was added to each block, block position being randomized. In a catch trial, the letter was not an *H* or *S*, but an *X*, and subjects were instructed to withhold the manual response, whatever the color of the stimulus, "if an *X* is presented". In group *SP* five catch trials were added to each block, each one signaled by a randomly-drawn member of the letter pool *A, B, C, D, W, X, Y*, and *Z*, and subjects were instructed to omit the manual response "if neither an H nor an *S* is presented". In group *RP* there were four additional trials per block in which the stimulus (i.e., the red or green *H* or *S*) was

surrounded by a white frame, in which case subjects were to vocally name the color of the stimulus letter instead of making the manual response. In group *SRP* there were eight additional trials per block, four catch trials construed as in group *SP* and four frame trials as used in group *RP*.

Results

Data from the dual-task phase were treated as in Experiment 1. In the single-task phase, premature responses and response omissions were excluded as well. Responses were correctly omitted in 4.1 of the 5 no go trials in group *NP*, in 23.4 of 25 trials in group *SP*, and in 18.5 of 20 trials in group *SRP*. Mean RTs for valid trials and PEs are given in Table 2.

The analysis of RTs from the two tasks in the first part of the experiment produced an effect of response, $F(1,76) = 1338.59$, and compatibility, $F(1,76) = 80.10$, and an interaction between these two factors, $F(1,76) = 20.75$. The latter indicated that compatibility effects were larger in secondary versus primary responses. However, the opposite was true in PEs, where the same effects were obtained: response, $F(1,76) = 8.86$, compatibility, $F(1,76) = 35.38$, and a response-by-compatibility interaction, $F(1,76) = 16.80$. This pattern suggests a speed-accuracy trade-off, in the sense that compatibility affected the primary task more in terms of errors than RTs but the secondary task more in terms of RTs than errors.

A comparison of the manual RTs across experimental halves yielded main effects of half, $F(1,76) = 126.52$, and compatibility, $F(1,76) = 67.05$, and significant interactions of half by group, $F(3,76) = 8.97$, and half by compatibility, $F(1,76) = 39.25$. The first interaction was due to the fact that practice reduced RTs rather strongly in groups *NP* and *RP* (226 and 272 ms, respectively) but less so in groups *SP* and *SRP* (106 and 88 ms). The latter interaction indicated a substantial reduction of the compatibility effect in the second half to 25–50% of its original size (see Table 2). However, a separate ANOVA

on manual RTs from the second half confirmed that the compatibility effect was still reliable and, importantly, unaffected by the group-specific manipulations. That is, the compatibility main effect was significant, $F(1,76) = 24.24$, while the interaction with group was not ($p > .8$). (Group had also no effect in an additional ANOVA of the changes in compatibility effect from the first to the second half, expressed as proportion of the effect size in the first half, $p > .6$.) The comparison of manual PEs across halves yielded a similar picture: Compatibility produced a main effect, $F(1,76) = 33.82$, and a close-to-significant interaction with half, $F(3,76) = 3.80$; $p < .06$, the latter indicating that the compatibility effect was smaller in the second than the first half. Again, however, the effect was still reliable in the second half, $F(3,76) = 14.54$, and independent from group ($p > .17$).

For the analysis of practice effects, manual RTs were grouped into blocks of 20 trials, separately for the four groups, the two experimental halves, and the two compatibility conditions. The first half produced a compatibility main effect, $F(1,76) = 77.27$, that was modified by block, $F(4,304) = 3.28$. As shown in Figure 2, the compatibility effect increased with increasing practice. The second half yielded a main effect of compatibility, $F(1,76) = 24.24$, and block, $F(4,304) = 27.17$, an interaction of block and group, $F(12,304) = 1.87$, and an interaction of compatibility and block, $F(4,304) = 3.07$. This latter interaction was due to that, on average, the compatibility effect was particularly pronounced in the first block of the second phase. However, separate comparisons confirmed that the effect was highly significant in all other blocks as well (p values $< .01$).

Discussion

Apart from once more replicating the backward-compatibility effect in its first part, Experiment 3 extends our previous results in two respects. First, significant

Table 2 Mean Reaction Time (RT, in ms) and Percentage of Errors (PE) for Manual (Primary) and Vocal (Secondary) responses in Experiments 3 (for the four Groups) and 4 according to half of session and compatibility between primary stimulus and secondary (Vocal) response. RT compatibility-effect Sizes (Δ, in ms) for each half in far right column, respectively. No secondary responses were performed in the second half of Experiment 3 in Groups NP and SP. (NP = No Preparation; SP = Stimulus Preparation; RP = Response Preparation; SRP = Stimulus/Response Preparation)

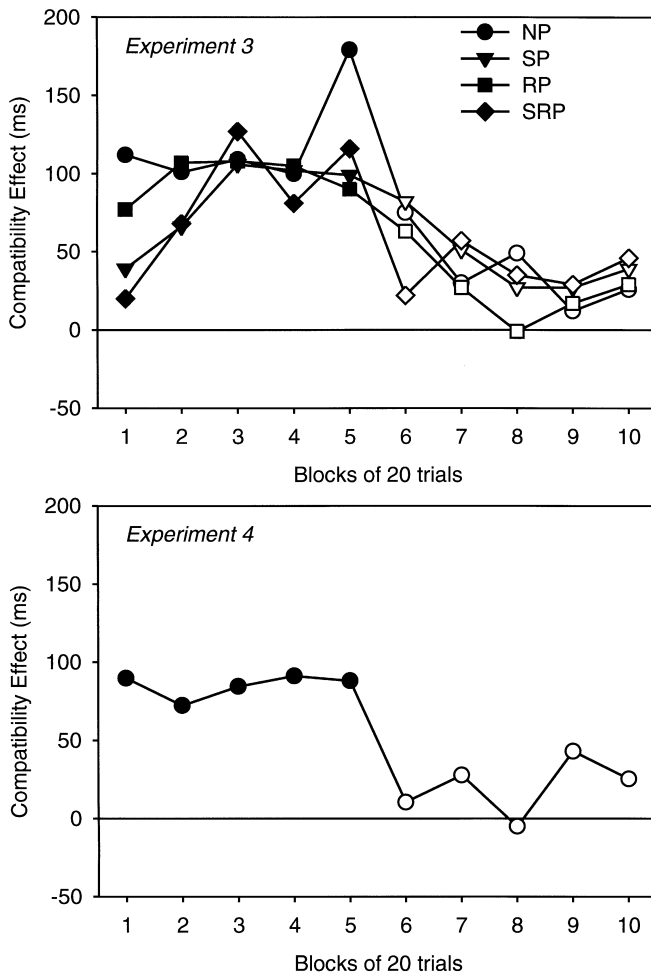|		First Half					Second Half				
		Compatible		Incompatible			Compatible		Incompatible		
Experiment	Response	RT	PE	RT	PE	Δ	RT	PE	RT	PE	Δ
3 (NP)	Primary	736	1.9	857	6.4	120	551	2.1	590	5.3	38
	Secondary	1304	3.1	1452	3.8	147					
3 (SP)	Primary	749	2.6	832	8.2	83	662	3.6	707	8.3	45
	Secondary	1350	2.0	1453	4.0	103					
3 (RP)	Primary	729	2.3	826	5.5	97	492	2.9	519	3.5	27
	Secondary	1302	1.7	1438	2.6	136	908	0.4	897	4.4	−11
3 (SRP)	Primary	700	2.1	783	5.6	84	634	3.7	672	5.6	38
	Secondary	1208	1.9	1337	3.9	129	1000	0.3	1047	0.7	47
4	Primary	743	2.5	829	6.9	85	780	2.2	801	2.7	20
	Secondary	1355	3.7	1445	7.1	90	1342	5.1	1366	3.9	24

Fig. 2 Compatibility effects (RTincompatible – RT$_{compatible}$) in primary, manual tasks of Experiments 3 and 4 as a function of 20-trial block (valid trials only) and preparation condition (in Experiment 3: NP = No Preparation; SP = Stimulus Preparation; RP = Response Preparation; SRP = Stimulus/Response Preparation). Black symbols represent blocks in first phase, white symbols represent blocks in the second phase, i.e., after the secondary task was dropped (in Experiment 3) or the S2-R2 mapping was switched (in Experiment 4)

compatibility effects were obtained in the second part as well, that is, after the subjects stopped performing the secondary task. This demonstrates that stopping to intend (or intending to stop) performing a task does not prevent S-R translation from being performed, which suggests that S-R translation satisfies even the most restrictive criteria of automaticity (intentionality criterion; see Neumann, 1984). The observation that compatibility effects were stronger immediately after stopping the secondary task is likely to indicate some kind of task inertia in the sense of Allport, Styles, and Hsieh (1994). That is, subjects seem to need some trials to "get out" the previous task set and to concentrate on the manual task. If so, the compatibility effects observed in the first block of the second phase are likely to overestimate the true effects of automatic S-R translation. Yet, reliable effects were obtained until the end of

the experiment, hence, after 100 trials, which should be enough to get into the new task. Accordingly, the effects observed in the later blocks can safely be taken to represent relatively pure measures of automatic translation. This is also suggested by the second important finding, namely the lack of any relationship between compatibility effects and preparation. Indeed, the four groups produced pretty much the same effect patterns, demonstrating that applying previously acquired S-R rules does not seem to depend on the readiness to process their input or output terms, i.e., the stimulus and response codes associated through these rules.

Experiment 4

The observation of compatibility effects in the second half of Experiment 3 shows that, once implemented or acquired, even arbitrary S-R associations outlive the task context they were created for. It also suggests that the application of those associations does not (directly) depend on task intentions and specific preparation processes. However, these conclusions are based on the assumption that our subjects strictly followed the instructions and carried out only those cognitive processes that were absolutely necessary to perform the task at hand. This assumption is certainly reasonable to make, but it is possible that we were not entirely successful to motivate our subjects to actively disengage from the secondary task in the second section of Experiment 3 and, hence, from translating S2 into R2. Alternatively, at least some subjects may not have been able to successfully disengage from the secondary task, or may have simply forgotten to do so. Indeed, although the RT level in the primary task did drop significantly in the second section, it was still rather high, certainly higher than one would expect from a true single-task condition. True, this may reflect the very existence of just acquired automatic S2-R2 translation, but we can not exclude other reasons.

To counter these kinds of possible problems we carried out Experiment 4, where we gave subjects very good reasons to forget about possible acquired S2-R2 associations. They started with exactly the same basic task as in Experiments 1–3. However, in the second half of the session the S-R mapping for the secondary, vocal task was switched, so that the same stimuli (i.e., the letters) now signaled the opposite secondary response than in the first half. The question was whether and how quickly this switch would affect automatic S2-R2 translation as indexed by backward compatibility. According to the transient-link model, one would have expected that the backward-compatibility effect immediately follows the changing mapping rule, so that the sizes of the effect in the first and second half should be comparable. According to the permanent-link model, however, the relevant S2-R2 associations would need to be changed, that is, the new associations would need to be formed and strengthened, and the old associations to be

unlearned or overruled by the new ones. Consequently, the size of the compatibility effect should shrink from the first to the second half.

Method

Twenty-four adults were paid to participate in a single session of about 30 min. They fulfilled the same criteria as in Experiment 1. Apparatus, stimuli, and procedure were as in the first section of Experiment 3. A session consisted of two sections, each comprising five 20-trial blocks composed as in Experiment 3. Primary responses were again assigned to stimulus color, secondary responses to letter identity, and the possible mapping combinations were counterbalanced across participants. However, the S2-R2 mapping was switched in the second section, i.e., subjects who in the first section responded to the letter *H* by saying "red" (or the German equivalent) and to *S* by saying "green" were in the second section responding to *H* by saying "green" and to *S* by saying "red", and vice versa.

Results

After treating the data as before, mean RTs and PEs for primary and secondary responses were computed as a function of half of experiment and (current) compatibility between R2 and S1 (see Table 2). The ANOVA of RTs yielded main effects of response, $F(1,23) = 568.89$, and compatibility, $F(1,23) = 19.65$, which were modified by interactions of half with response, $F(1,23) = 5.20$, and of half with compatibility, $F(1,23) = 5.33$. RTs in the secondary task decreased from the first to the second half (1400 vs. 1354 ms), whereas primary-task RT remained constant (786 vs. 790 ms). More important, the compatibility effect decreased dramatically from 88 ms in the first half to 22 ms in the second, the latter being no longer reliable in a separate test. To check for possible learning effects within experimental halves, especially directly following mapping switch, the manual RTs were grouped into blocks of 20 trials and analyzed separately for the two experimental halves, as a function of compatibility and block. However, as evident from Figure 2, there was no evidence of systematic learning effects, i.e., compatibility did not reliably vary with block before or after the mapping was switched. The ANOVA of PEs yielded main effects of response, $F(1,23) = 7.03$, and compatibility, $F(1,23) = 6.22$, and an interaction between compatibility and half, $F(1,23) = 7.05$. Errors in the secondary task were more frequent than in the primary task (5.0% vs. 3.6%). More important, compatible conditions produced less errors than incompatible conditions; yet, this effect was reliable in the first half only (3.1% vs. 7.0%) but not in the second half (3.6% vs. 3.2%).

Discussion

Switching the mapping of the secondary task had a strong impact on backward-compatibility effects and actually eliminated them. This means that backward effects are not a direct function of the currently valid S-R mapping rule, as the transient-link model would have assumed, but seem to emerge with practice over time. Apparently, then, automatic translation is mediated by acquired, rather stable, direct connections between stimulus and response codes that effectively circumvent short-term memory systems. Once implemented and sufficiently strengthened, these connections work against linking the same codes in a different way, i.e., they compete with new associations. In fact, in this experiment the effects of the old associations canceled out those of the new ones completely, so that no backward effect was obtained. Of course, with extensive practice the new associations are likely to eventually prevail and, thus, produce backward effects that reflect the post-switch mapping. However, it is also clear from our data that the amount of practice required would need to be considerably greater than that sufficient for acquiring the old associations.

General Discussion

The aim of this study was to understand and characterize the mechanism underlying automatic S-R translation as evidenced by backward-compatibility effects under dual-task conditions. In particular, we tested and compared two translation models. According to the transient-link model, S-R translation is always and exclusively mediated by S-R rules temporarily implemented in working memory. Once a rule is implemented, it can be accessed by any kind of stimulus information, which then leads to the activation of the response specified in the rule. There are, however, several aspects of our data that are inconsistent with such a model. First, backward-compatibility effects were unaffected by increasing the load of working memory, either by an additional memory task (Experiment 1) or by increasing the number of S-R rules in the primary task (Experiment 2). Such a finding does not only contradict the transient-link model under test, it is inconsistent with any model that assumes S-R translation (not response selection) to be inevitably mediated by limited-capacity systems, processes, or stages, such as that of Welford (1952) and followers. Moreover, our findings underline the necessity to conceptually distinguish between S-R translation and response selection– possible capacity limitations of the latter in no way imply limitations in the former. Second, backward-compatibility effects were not restricted to the time period in which the rules they reflect were needed. Experiment 3 provided evidence that the rules for the secondary task were still applied 100 trials after performing this task was stopped, and Experiment 4 indicated that implementing new rules did not lead to reliable backward effects if these rules contradicted those applied previously. That is, discontinuing a task is not associated with the immediate disabling of the corresponding rules. In fact, the outcome of Experiment 4

suggests that old rules are not disabled even if this would be in the interest of the performing individual. These observations are not only inconsistent with the transient-link model and other models assuming capacity limitations in S-R translation; they also provide problems for unlimited-capacity models, such as Meyer and Kieras' (1997) EPIC model. According to this model, S-R translation is mediated by production rules stored in a working memory of unlimited capacity. Although such a model can nicely deal with the parallel application of primary- and secondary-task rules, to account for the present findings it would need to assume that people are unwilling and even unable to eliminate no longer useful secondary-task rules from working memory within 100 trials. This does not seem to fit with the high degree of adaptivity and flexibility of executive control EPIC explicitly assumes.

According to the theoretical alternative we considered, the permanent-link model, repeatedly performing a particular response to a particular stimulus leads to the formation of some kind of direct association between the corresponding stimulus and response codes. Automatic S-R translation proceeds along these direct links and is, therefore, independent of the task, the context, and the available working-memory capacity once the associations are sufficiently strengthened. This approach is consistent with the outcome of Experiments 1 and 2, as this showed no impact of increasing the demands on working memory on backward-compatibility effects. It is also in agreement with our observations in Experiments 3 and 4 that the S-R links mediating automatic translation seem to outlive the task that led to their formation. Converging evidence for such after-effects of S-R translation comes from studies of task-switching performance. For instance, Allport and Wylie (2000) presented subjects with Stroop-type word-color pairs (e.g., the word "green" written in red ink) and had them switch between naming the word and the color. Performance was worse in switch than repetition trials– a common observation (e.g., Allport et al. 1994; Rogers & Monsell, 1995). However, switching costs were substantially larger if the relevant stimulus had already appeared as distractor under the other task than if it did not. Likewise, Waszak, Hommel, and Allport (2002) found that switching from naming the picture to naming the word of incongruent word-picture pairs took about twice as long if the present picture component had been previously named at least once, even if that happened more than 100 trials earlier. Apparently, then, even a single response to a stimulus is sufficient to create a direct, stable S-R association, which is automatically retrieved if the same stimulus is encountered again (Hommel, 1998b; Hommel, Pösse, & Waszak, 2000; Logan, 1988). Our present findings fit nicely into this picture and extend it in showing evidence of automatic retrieval of S-R associations even in "non-switch" tasks and even after the corresponding task is no longer carried out.

An objection one may have with respect to the present findings relates to the wide range of compatibility-effect sizes we obtained, from 180 ms in Experiment 1, to 27 ms in the *RP* group of Experiment 3. What factors are responsible for these differences? In our view, the answer to this question may be relatively straightforward. In the original study, Hommel (1998a) performed analyses of the RT distribution in backward-compatibility tasks and observed that the compatibility effect on the primary task was consistently positively correlated with RT in the primary task. This is what one would expect: The longer it takes participants to perform the primary task, the more time is available for processing the secondary stimulus, activating the secondary response, and then backward-priming the compatible stimulus code in the primary task. However, this means that any modification that speeds up performance in the primary task is likely to work against the compatibility effect and to reduce its size. If so, one would expect a decrease of effect size under single-task conditions to the degree that this is accompanied by a reduction in manual RT. Indeed, single-task RTs were markedly reduced, suggesting that manual reactions may have often been too quick to be affected by backward priming. Moreover, we carried out additional analyses revealing substantial, positive correlations between (a) overall size of the manual compatibility effect and average manual RT in the particular task, (b) individual effect size and individual overall mean, and (c) individual effect size and quintile of the individual distribution of manual RTs. Hence, whatever the measure, the compatibility effect is the bigger the more time is available to apply the mapping rules of the present or previous secondary task.

All in all, the present findings point to the existence of automatic, capacity-free translation of stimulus features into arbitrarily mapped responses, mediated by relatively quickly emerging, robust S-R associations. Of course, demonstrating the existence of automatic translation by no means rules out important roles of other, non-automatic, capacity-limited processes concerned with stimulus-related responding. Authors from Ach (1910) to Logan (1988) have proposed that stimulus information is processed concurrently via slow, serial, and highly controlled pathways *and* quick, parallel, and automatic pathways, with the latter becoming more efficient and dominant over practice. That is, even though in our experiments secondary stimuli were apparently translated into secondary responses in parallel to primary-task S-R translation, this does not preclude contributions from other, more controlled and perhaps serial-working response-selection processes. Indeed, several aspects of our findings point to such processes: Clearly, the vocal RTs in all four experiments are way too long to believe that the two tasks were entirely carried out in parallel. True, our design required subjects to deliver responses in succession and thereby enforced seriality. Yet, the imposed minimal RT delay of 50 ms seems too short to explain RTs in a rather simple verbal task of more than a second. Moreover, to have some impact on processing compatible or incompatible stimuli in the primary task, secondary responses must have been

activated (to some degree) at about 100–200 ms, or even longer, before the primary response was executed. In view of primary RTs around 700 ms, this means that secondary responses were activated no later than about 500 ms after stimulus presentation. However, if such R2 priming would have been immediately followed by R2 selection, it would be difficult to explain why carrying out the secondary response took another 500–1000 ms. In other words, automatic S-R translation as indicated by backward priming is unlikely to directly subserve eventual response selection– which is also indicated by the fact that compatibility effects on the secondary task commonly reflect little more than mere propagation of the effect from the primary task. Why people do not use response information delivered by automatic S-R translation is still an open question. It is possible that they deliberately wait until controlled, serial S-R translation or response re-checking processes are completed, just to make sure that the correct mapping rule is applied and/or that the responses are carried out in the right order (Meyer & Kieras, 1997; Ruthruff et al., 2001). Alternatively, automatic S-R translation and response-code activation may only be the first step of action planning, which is then to be followed by the (presumably serial) integration of those response codes that belong to the same response (Stoet & Hommel, 1999). Hence, the present findings provide evidence that some S-R translation proceeds automatically and independently of cognitive capacity, preparation and task intentions, not that all aspects of S-R translation or response selection are automatic.

References

Ach, N. (1910). *Über den Willensakt und das Temperament* [On will and temperament]. Leipzig: Quelle & Meyer.

Allport, A., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Allport, D. A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35–70). Cambridge, MA: MIT Press.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation, 8*, 47–89.

Berlyne, D. E. (1957). Uncertainty and conflict: A point of contact between information-theory and behavior-theory concepts. *Psychological Review, 64*, 329–339.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 965–980.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*, 249–263.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1058–1079.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (ed.), *Categories of human learning* (pp. 243–285). New York: Academic Press.

Hommel, B. (1998a). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1368–1384.

Hommel, B. (1998b). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183–216.

Hommel, B. (2000a). The prepared reflex: Automaticity and control in stimulus-response translation. In S. Monsell & J. Driver (eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 247–273). Cambridge, MA: MIT Press.

Hommel, B. (2000b). Intentional control of automatic stimulus-response translation. In Y. Rossetti & A. Revonsuo (eds.), *Interaction between dissociable conscious and nonconscious processes* (pp. 223–244). Amsterdam: John Benjamins Publishing Company.

Hommel, B., Pösse, B., & Waszak, F. (2000). Contextualization in perception and action. *Psychologica Belgica, 40*, 227–245.

Jolicœur, P., Tombu, M., Oriet, C., & Stevanovski, B. (2002). From perception to action: Making the connection. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention & Performance XIX.* (pp. 558–586) Oxford: Oxford University Press.

Jonides, J., & Smith, E. E. (1997). In M. D. Rugg (ed.), *Cognitive neuroscience* (pp. 243–276). Hove: Psychology Press.

Keele, S. W., & Boies, S. J. (1973). Processing demands of sequential information. *Memory & Cognition, 1*, 85–90.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility– A model and taxonomy. *Psychological Review, 97*, 253–270.

Logan, G. D. (1978). Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107*, 32–63.

Logan, G. D. (1979). On the use of the concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 189–207.

Logan, G. D. (1980). Short-term memory demands of reaction-time tasks that differ in complexity. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 375–389.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.

Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1072–1090.

Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2*, 174–207.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 471–484.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104*, 3–65.

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In W. Prinz & A. F.

Sanders (Eds.), *Cognition and motor processes* (pp. 255–293). Berlin: Springer.

Pashler, H. (1993). Doing two things at the same time. *American Scientist, 81*, 48–55.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology, 41 A,* 19–45.

Prinz, W. (1984). Modes of linkage between perception and action. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 185–193). Berlin: Springer.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Berlin: Springer.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.

Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 3–21.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor, & T. G. Reeve (Eds.), *Stimulus-response compatibility* (pp. 31–86). Amsterdam: North-Holland.

Smith, M. C. (1967). The psychological refractory period as a function of performance of a first response. *Quarterly Journal of Experimental Psychology, 19*, 350–352.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1625–1640.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*, 212–222.

Waszak, F., Hommel, B., & Allport, A. (2002). Task-switching and long-term priming: Role of episodic stimulus-response bindings in task-shift costs. *Cognitive Psychology* (in press)

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance – a review and a theory. *British Journal of Psychology, 43*, 2–19.