HHS Public Access

Author manuscript

Nat Rev Neurosci. Author manuscript; available in PMC 2016 November 29.

Published in final edited form as:

Nat Rev Neurosci. 2016 February ; 17(2): 77–88. doi:10.1038/nrn.2015.18.

The right time to learn: mechanisms and optimization of spaced learning

Paul Smolen, **Yili Zhang**, and **John H. Byrne**

Department of Neurobiology and Anatomy, W. M. Keck Center for the Neurobiology of Learning and Memory, The University of Texas Medical School at Houston, P.O. BOX 20708, Houston, Texas 77030, USA

Abstract

For many types of learning, spaced training, which involves repeated long inter-trial intervals, leads to more robust memory formation than does massed training, which involves short or no intervals. Several cognitive theories have been proposed to explain this superiority, but only recently have data begun to delineate the underlying cellular and molecular mechanisms of spaced training, and we review these theories and data here. Computational models of the implicated signalling cascades have predicted that spaced training with irregular inter-trial intervals can enhance learning. This strategy of using models to predict optimal spaced training protocols, combined with pharmacotherapy, suggests novel ways to rescue impaired synaptic plasticity and learning.

Repetitive training helps to form a long-term memory. Training or learning that includes long intervals between training sessions is termed spaced training or spaced learning. Such training has been known since the seminal work of Ebbinghaus to be superior to training that includes short inter-trial intervals (massed training or massed learning) in terms of its ability to promote memory formation. Ebbinghaus stated: “with any considerable number of repetitions a suitable distribution of them over a space of time is decidedly more advantageous than the massing of them at a single time” (REF. 1). His studies were based on the self-testing of acquired memory for lists of syllables, but the superiority of spaced training has now been established for many additional forms of human learning. For example, spaced learning is more effective than massed learning for facts, concepts and lists[2–4], skill learning and motor learning[5,6], in classroom education (including science learning and vocabulary learning)[7–9], and in generalization of conceptual knowledge in children[10]. Spaced training also leads to improved memory in invertebrates, such as the mollusk *Aplysia californica*[11–14], *Drosophila melanogaster*[15,16] and bees[17], and in rodents[18,19] and non-human primates[20,21]. Memory extinction is commonly considered to involve the formation of a new memory, and in rat fear conditioning spaced extinction trials are more effective than massed trials at establishing new memories[22].

Correspondence to J.H.B. John.H.Byrne@uth.tmc.edu.

Competing interests statement

The authors declare no competing interests.

Although it has been established that spaced training is superior to massed training in terms of inducing memory formation, key questions remain. What are the mechanisms underlying this superiority? Is it possible to use this mechanistic information to determine the optimal intervals between learning trials? If so, are fixed, expanding or irregularly spaced intervals optimal? Another key question is whether an understanding of the mechanisms for optimal intervals can provide insights into the design of pharmacological approaches for memory enhancement. Computational models based on such a mechanistic understanding may be able to predict more complex approaches to memory improvement in which the application of multiple drugs, or combinations of drugs and training protocols, can enhance memory or treat deficits in learning and memory.

In this Review, we describe how new insights from molecular studies may help to explain the effectiveness of spaced training, and how the molecular findings relate to the traditional learning theories that aim to account for this effectiveness. We also review how models of signalling pathways that are involved in synaptic plasticity can suggest, and experiments empirically validate, training protocols that improve learning and that rescue plasticity impaired by deficits of key molecular components. Finally, we discuss recent models that have suggested combined-drug therapies that may further enhance some forms of learning and that may have synergistic effects with optimized spaced learning on memory formation.

Traditional learning theories

We briefly summarize three of the well-known cognitive theories that have been proposed to explain the superiority of spaced training over massed training: encoding variability theory, study-phase retrieval theory and deficient-processing theory.

Encoding variability theory[23–25] posits that repeated stimulus presentations or learning trials are more likely to occur in multiple contexts if they are spaced further apart in time and that a memory trace for repeated trials therefore includes elements of each of these contexts. Thus, spaced training would tend to bind together more contexts and hence form a more robust memory, as a greater number of testing contexts could elicit retrieval of the memory.

Study-phase retrieval theory[26–29] posits that spaced stimulus presentations or learning trials are more effective than massed trials for memory reinforcement because each spaced trial elicits retrieval and reactivation of a memory trace that was formed by the preceding trial. By contrast, with short massed trials, the preceding memory trace is still active, so it is not retrieved or reactivated and therefore the memory cannot be reinforced. Study-phase retrieval theory also accounts for a decline in learning in trials with excessively long intervals because in those cases the preceding memory trace can no longer be retrieved. A recent variant, retrieved context theory, also incorporates elements of encoding variability theory and has succeeded in predicting the results of subsequently performed spaced learning experiments in humans[30].

Deficient-processing theory posits that spaced training forms a stronger memory than does massed training because, in the latter, some processes that are necessary to form memories are not effectively executed. The reasoning here becomes clearer by examining variants of

this theory that specify the nature of the deficient process. One variant posits that excess habituation during massed trials prevents effective reinforcement of memory traces[31], whereas others posit that there is a failure to consolidate a memory (known as consolidation theory)[32,33], a lack of voluntary attention to massed presentations[31,34], or a lack of cognitive rehearsals or memory reactivations within the short intervals that are characteristic of massed training[27,35].

Consolidation occurs as a memory trace becomes more fixed and stable with time after training[2,36]. Thus, consolidation theory[37–39] posits that a long-term memory trace is more efficiently stabilized or strengthened by spaced trials. The lack of cognitive rehearsals variant of deficient-processing theory might also be considered a more specific form of consolidation theory, because it assumes that a minimum number of rehearsals, or autonomous reactivations, are required to consolidate a memory trace. Variants of deficient-processing theory and relevant experiments are discussed in more detail in REFS 28,40,41. Below, we focus substantially on consolidation theory because, of all the traditional learning theories, it seems to be most closely aligned with our current understanding of the cellular and molecular mechanisms of memory.

Landauer[37] was one of the first researchers to develop a conceptual model of the ways in which consolidation principles could explain the effectiveness of spaced training. Although the model was originally developed to explain the effects of short spacing intervals on memory formation, it can readily be generalized for the effects of arbitrarily long intervals (FIG. 1). The model is based on two assumptions. First, the state of a neural circuit following the first learning trial is such that a second reinforcing trial soon after will not markedly increase the consolidation of the learning trace resulting from the first trial (FIG. 1a). Thus, in massed training, overlap between traces, which may cause saturation of an unspecified molecular mechanism, diminishes the summed impact of the traces on the consolidation of memory. Only when the effects of the first trial decay can the effects of a second trial be fully expressed (FIG. 1b,c), leading to greater potential consolidation of a memory in spaced training than in massed training (greater net gain; see FIG. 1d). The second assumption is that the probability that the second trial can successfully reinforce the first trial declines with time (FIG. 1e). Actual consolidation is the product of these two assumptions, yielding a prediction of an optimal interval for spaced learning (FIG. 1f).

Peterson[38] described a similar model that focused on the dynamics of verbal learning. Furthermore, Wickelgren[39] extended consolidation theory by positing that the resistance of a memory trace to decay increases with the age of the trace over the total duration of a spaced learning protocol. Thus, a trace would become not only strong but also highly resistant to decay following spaced trials.

Molecular traces of time

Substantial progress has been made in understanding the molecular mechanisms of memory. Given this progress, in this section we focus on potential molecular mechanisms of the spacing effect on long-term memory formation.

There is now agreement that learning is implemented, at least in part, by changes in synaptic strength (synaptic plasticity). For example, fear-conditioned memories can be alternately erased and reinstated by long-term depression (LTD) and long-term potentiation (LTP), respectively, of a defined synaptic pathway[42]. Thus, the molecular processes that are essential for spaced learning might reinforce extant LTP.

Reliable correlates of LTP are the remodelling and enlargement of postsynaptic dendritic spines, which are small protrusions that are associated with most excitatory synapses[43]. Thus, studying the differential dynamics of dendritic spine remodelling following massed versus spaced stimuli is likely to provide insight into processes underlying the effectiveness of spaced training. Studies using rat hippocampal slices found that LTP induced by multiple trains of theta-burst stimuli was accompanied by extensive remodelling of synaptic ultrastructures[44,45] and that subsequent spaced trains of theta-burst stimuli, with intervals of 60 minutes or more between the trains, were needed for optimal reinforcement of LTP[46]. Stimulated dendritic spines were remodelled over a period of more than 1 hour, leading to enlargement of the existing functional postsynaptic density[45] and the presynaptic active zone[44]. The resulting increase in the numbers of AMPA-type and NMDA-type glutamate receptors at the synapse correlated with the magnitude of LTP.

Two hypotheses that involve spine remodelling have been put forward to explain the greater efficacy of spaced trials over massed trials in memory formation. These hypotheses have a common theme, which is that the learning process includes a refractory period during which the second of two closely spaced stimuli would be ineffective in enhancing the effects of the first (FIG. 2a). One hypothesis is that spaced but not massed repetitions of a stimulus allow the refractory period to be overcome and lead to repeated enlargement of a set of spines and strengthening of the synaptic connections mediated by these spines[47] (FIG. 2b). A second, not mutually exclusive, hypothesis[47,48] is that molecular processes enable later spaced stimuli to induce LTP at spines that do not undergo initial enlargement. In this case, spaced, but not massed, inter-trial intervals would allow for a molecular process termed ‘priming’ to be completed at these additional spines. After being primed, these spines would be strengthened by subsequent stimuli and incorporated into the memory trace (FIG. 2c). Currently, the molecular components of such a priming process are not known.

Through the use of Schaffer–commissural projections in rat hippocampal slices, two studies[47,48] have characterized the recruitment of additional synaptic contacts with the application of spaced stimuli. Theta-burst stimuli applied at intervals of 10 or 40 minutes did not cumulatively increase LTP. However, for longer intervals (60 or 90 minutes), a cumulative increase in LTP was observed over three bursts of stimulation. Each theta-burst stimulus led to actin filament polymerization in spines, which is known to be important for the stabilization of LTP[49]. The second theta-burst stimulus yielded polymerization in spines that were not apparently affected by the first stimulus, if the second followed the first by 60 or 90 minutes. These data do not suggest that successive theta-burst stimuli further strengthen the efficacy of the same spines. Instead, they suggest that the first theta-burst stimulus initiates priming at all synaptic contacts of the stimulated afferents but only initiates consolidation and strengthening at a subset of contacts. Spines that undergo priming but not consolidation exhibit a refractory period of ~60 minutes, suggesting that priming

takes time to complete (FIG. 2a). If the second theta-burst stimulus is applied after the refractory period, some or all of the primed spines undergo consolidation. These data are consistent with the second hypothesis presented in the preceding paragraph, because the first theta-burst stimulus appears to enlarge and strengthen some spines but, at others, it only initiates priming. These primed spines can then be strengthened by the second theta-burst stimulus.

The dynamic properties of transcription factors and their interactions could also account for the superior efficacy of spaced training. LTP that persists for several hours or more requires translation and transcription[50,51], which is reliant on key transcription factors such as cyclic AMP-responsive element (CRE)-binding protein (CREB)[52]. Spaced training may be more effective, in part, because it may allow sufficient time for transcription factors such as CREB to be activated, bind to promoters and induce a round of transcription for the consolidation of LTP[53] or for long-term facilitation (LTF) of synapses[54]. In massed training, the trials would come too close together to initiate separate rounds of transcription. Indeed, in co-cultures of sensory and motor neurons from *A. californica*, five spaced applications of 5-hydroxytryptamine (5-HT; also known as serotonin), each lasting 5 minutes with an inter-stimulus interval of 20 minutes (an analogue of spaced training), robustly elicit LTF that lasts for more than 24 hours[14], whereas 5-HT applied continuously over 25 minutes (an analogue of massed training) fails to yield reliable LTF.

In these sensory neurons, levels of the transcription activator CREB1 are elevated for at least 24 hours after the spaced 5-HT treatments[54,55]. This prolonged elevation of CREB1 levels is due to a positive feedback loop in which this protein, by binding to a CRE regulatory element near *creb1*, increases the expression of *creb1* (REFS 54,55) and other genes that are upregulated by CREB1. In addition, in these sensory neurons, the level of the transcription repressor CREB2 shows a late drop at ~12 hours after treatment[56]. This drop in the level of CREB2, coupled with the rise in the level of CREB1, plausibly corresponds to an increased potential for gene induction. Thus, an additional 5-HT pulse near 12 hours after treatment might optimally reinforce LTF.

LTF at these sensorimotor synapses is associated with a simple form of learning, long-term sensitization (LTS) of withdrawal reflexes. *In vivo*, four spaced electrical stimuli (with 30 minutes intervals between the stimuli) yielded LTS that lasted for more than 24 hours, with weak residual LTS being detectable at 4 days post-training, and repetition of this spaced protocol once per day for 4 days yielded much stronger LTS that lasted for more than 1 week[13,57,58]. These data suggest that, in this system, the dynamics of transcription activation and gene expression have slow components that can summate over multiple days, yielding long-lasting memory.

Recent data also illustrate that, in the hippocampus, CREB and CCAAT enhancer-binding protein (C/EBP), another transcription factor that is important for LTP, can remain active for many hours after learning. Following inhibitory avoidance training in rats, late peaks in brain-derived neurotrophic factor (BDNF) expression and in C/EBP expression occur at ~12 hours post-training, and inhibiting BDNF action at this time blocks memory maintenance[59]. These BDNF dynamics result from a positive feedback loop in which *C/ebp* induction leads

to *Bdnf* upregulation, with the resulting increase in BDNF levels further activating the C/EBP signalling pathway[60]. Although this slow feedback loop was activated by single-trial training rather than spaced training, it would be of interest to model these dynamics, and to examine whether an additional spaced trial at ~12 hours post-training, leading to a second induction of *C/ebp* at the time of elevated C/EBP levels, might optimally reinforce learning. A second prediction would be that massed stimuli are less effective if repeated at an interval too brief to allow the transcription regulation, and thus *Bdnf* expression, that is necessary to activate this feedback loop. Insights that can be obtained from computational models of learning are discussed later in the article.

On a shorter timescale, the dynamics of second messengers, kinases and phosphatases may contribute to the superiority of spaced training. One study in mice[61] found marked phosphorylation and activation of CREB in the hippocampus and the cortex when object recognition trials were separated by an interval of 15 minutes but not by an interval of 5 minutes. Protein phosphatase 1 (PP1) appeared to be necessary for this spacing effect, because PP1 inhibition allowed the shorter interval to activate CREB. A study involving *A. californica* sensory neuron–motor neuron co-cultures[62] found that protein kinase C (PKC) is activated to a greater extent during a massed stimulus (continuous 5-HT application) than during a spaced stimulus (15-minute intervals between applications). It is known that PKC acts to downregulate protein kinase A (PKA) and that PKA activation is necessary for LTF; thus, these data delineate crosstalk between signalling pathways such that LTF is suppressed, in part, by stronger PKC activation during massed training.

Another study[16] characterized the dynamics of mitogen-activated protein kinase (MAPK) and of MAPK phosphatase in *D. melanogaster*. In an olfactory learning protocol, each spaced training trial generated a distinct wave of MAPK activity, whereas massed training trials were too close together to generate distinct waves. The authors therefore hypothesized that effective learning depended on the generation of distinct waves of MAPK activity.

Another phosphorylation-based mechanism has also been hypothesized to help to explain the efficacy of spaced intervals in *D. melanogaster*. Spaced (15-minute) intervals were more effective than massed (1-minute) intervals in inducing olfactory learning, even given the same total training time (and thus more massed presentations)[63]. Two isoforms of *D. melanogaster* CREB — dCREB2-a and dCREB2-r — can activate and repress transcription, respectively. The authors proposed[64] that the kinetics of the phosphorylation of these isoforms differed such that the kinase activation generated by less frequent, spaced trials was sufficient to phosphorylate and activate dCREB2-a, whereas dCREB2-r could only be effectively phosphorylated by massed trials. Thus, training involving spaced intervals could maximally activate transcription and possibly induce the formation of long-term memory by activating dCREB2-a but not the counteracting repressor dCREB2-r.

Computational simulations have supported the plausibility of this mechanism[65], but it has not been validated empirically. However, it appears to be likely that a similar type of mechanism that is based on competition between an activator of long-term memory formation and a repressor, with the repressor only activated at short intervals, might be

needed to explain any similar data in which massed training is less effective than spaced training even given equal total training times.

In experiments with *A. californica*, when two electric shocks were given to induce LTS, maximal LTS was produced when the inter-stimulus interval was 45 minutes. LTS was not produced with intervals of 15 or 60 minutes[66]. The 45-minute optimum was associated with activation of MAPK. Following either a single 5-HT pulse or a single electric shock, MAPK activation peaked at or near 45 minutes post-trial[12,66]; thus, a 45-minute interval might optimally reinforce the effects of MAPK. It is known that this delayed MAPK activation requires protein synthesis[12], although the upstream mechanisms underlying the dynamics of the peak in MAPK activity at ~45 minutes are not well understood. Nevertheless, the key finding from these studies is that delayed activation of MAPK is intimately associated with the effectiveness of spaced stimuli to induce long-term memory.

Similarly, training with intervals of 60 minutes, but not 20 or 120 minutes, enhanced object recognition learning in wild-type mice and in a mouse model of fragile X syndrome (fragile X mental retardation 1 (*Fmr1*)-knockout mice), at least partly by increasing synaptic activation of extracellular signal-regulated kinase 1 (ERK1; also known as MAPK3) and ERK2 (also known as MAPK1)[67]. This 60-minute interval was predicted to be optimal for learning because stimuli separated by 60 minutes had previously been found to enhance LTP in wild-type rodents[47]. Thus, in *A. californica*, *D. melanogaster* and mammals, MAPK activation appears to be a component of the molecular mechanism that underlies the spacing effect.

Some of these molecular mechanisms appear to fit with a theory in which spaced training sessions are effective because they reinforce the same memory trace or group of strengthened synapses. However, spaced stimuli might also reinforce memory by recruiting new synapses. ERK1 and ERK2 (ERK1/2) activation is needed for some forms of LTP[68], and one study[69] compared ERK1/2 activation in rat hippocampal pyramidal neurons following three spaced tetanic bursts (at 5-minute intervals) with that after three massed bursts (at 20-second intervals). About twice as many dendrites with active ERK1/2 were found following spaced bursts, suggesting that spaced trials may recruit additional synapses on different dendrites for LTP. Thus, a range of molecular and cellular mechanisms appears to contribute to the efficacy of spaced training, in parallel or in series.

An extremely broad range of inter-trial intervals, from seconds to days, has been used for spaced training (FIG. 3). For example, in honeybee olfactory learning, efficient spaced training can occur with intervals as short as 1 minute[17]. Such brief intervals might allow for the reinforcement of the activity of a short-lived second messenger such as cAMP that is produced by preceding trials. The dynamics of kinase activation constitute a second substrate of spacing effects. In *A. californica*, *D. melanogaster* and mammals, the data discussed above indicate that commonly reported intervals, ranging from ~5 minutes to 1 hour, may allow for the reinforcement of the activities of key kinases essential for LTP or LTF, and consolidate structural changes in dendritic spines.

It is plausible that the minimum inter-stimulus interval for effective learning, for a given protocol and system, corresponds to the interval that is necessary to allow each stimulus to contribute separately to a rate-limiting biochemical process. For example, for rapid honeybee olfactory learning with an effective interval of 1 minute, the rate-limiting process might be second messenger accumulation or rapid activation of a kinase. For even shorter intervals, the timescale of the rate-limiting process might be too long to permit each brief stimulus to contribute separately to the process — a group of closely spaced stimuli would instead tend to act as just a single stimulus. For intervals of 1 minute or more, each stimulus would be able to contribute a discrete increment to the rate-limiting process, allowing effective learning. For the spaced LTP protocol of Gall, Lynch and colleagues[47,48], an interval of 40–60 minutes is needed for successive theta-burst stimuli to further increase LTP. Here, the rate-limiting process would be different — plausibly slower activation of an unspecified kinase or other intracellular signalling event, with a time constant near the minimum effective interval of ~40 minutes. Stimuli at intervals much shorter than this would not be able to generate summation of the rate-limiting process and would therefore not cause additional LTP.

For other systems, a similar assumption may apply to the dynamics of transcription activation. For LTF and LTS in *A. californica*, transcription, as discussed above, may constitute a rate-limiting process that helps to determine the efficacy of spaced training. However, it is evident that even for systems such as honeybee olfactory learning that involve short, spaced intervals, effective long-term memory formation relies on the activation of transcription and translation, downstream of the intracellular signalling pathways that are activated by these intervals[17,70]. Reactivation of memory traces may constitute an additional temporal substrate that underlies the longest reported effective intervals, on the order of a week[71]. Such intervals are likely to reactivate and reinforce consolidated patterns of strengthened synapses that correspond to memory traces that are maintained by neuronal network activity[72]. Spaced learning with these long intervals would reactivate critical components at these synapses, and in particular reactivate NMDA receptors at these synapses. Studies using inducible and reversible NMDA receptor knockouts have demonstrated that such NMDA receptor reactivation, which may also in part result from spontaneous neuronal activity, is required to sustain remote memory storage[73,74]. Positive feedback loops that maintain key kinases and other molecules in persistently active states at strengthened synapses may also contribute to such long-term memory storage[75–79]. An important topic for future research will be to further investigate the molecular processes that support effective spaced learning in humans that involves inter-trial intervals of a day or more.

An implication of the work outlined above is that multiple temporal domains of spaced training may be engaged in spaced training (FIG. 3). Indeed, an effective protocol for LTS training in *A. californica* is the use of four trials with an inter-trial interval of 30 minutes, repeated four times with a 1-day inter-trial interval[13]. Thus, at least in some cases, there appears to be a hierarchy of temporal domains of training protocols, with briefer protocols embedded within longer ones.

The above considerations, and most empirical studies, are concerned with only typical, or minimum, inter-trial or inter-stimulus intervals for effective spaced learning or for the summation of LTP. Only a few studies have delineated, for any specific system (that is, a given species and stimulus protocol), both minimum and maximum effective intervals. One study[80] found that in a hippocampal slice preparation, 5–10-minute intervals between tetani were ideal for induction of LTP, and they produced similar levels of LTP, with longer or shorter intervals yielding both less LTP and less ERK1/2 activation. In *A. californica*, LTS was effectively induced by an interval of 45 minutes between electrical stimuli, but not by intervals of 15 minutes or 60 minutes[66]. As noted above, the authors of this study hypothesized that the coincidence of peak MAPK activation with the second trial was necessary for effective learning. In addition, 60-minute intervals were effective for forming object location memory in mice with three trials, but intervals of 20 minutes or 120 minutes were not[67].

Owing to the small number of such studies and the lack of sufficient characterization of the accompanying molecular processes, it is not yet possible to make detailed statements about the ways in which intracellular signalling pathways could cooperate to generate both minimum and maximum intervals. For maximum intervals, a reasonable qualitative assumption is that each trial or stimulus generates a separate, relatively short-lived biochemical trace and that, for effective spaced learning, these traces must overlap and summate, with the summed magnitude driving long-lasting synaptic potentiation. These dynamics would be analogous to the necessary overlap of traces in the conceptual model of Landauer (FIG. 1a–c). For intervals longer than the maximum, the individual biochemical traces would decay and not overlap.

Recent data and learning theories

Do the biochemical and morphological mechanisms that are proposed to contribute to the greater efficacy of spaced training align with traditional cognitive theories? At this point, much of the extant cellular data seem to be compatible with the deficient-processing theory, particularly two of its variants: the consolidation theory and the lack of cognitive rehearsals theory. In the consolidation theory, intervals between massed trials are proposed to be too short for the consolidation and consequent summation of memory traces that are engendered by successive trials. In the cognitive rehearsals theory, massed trials are proposed to lead to fewer cognitive rehearsals, or autonomous reactivations, than do spaced trials, and therefore less cumulative consolidation and persistence of a memory.

The required refractory period of ~1 hour between successive theta-burst stimuli to induce progressive increments in hippocampal LTP[46,47] may be in line with the first of these variants, which is that short intervals are insufficient for consolidation and consequent summation of memory traces. The refractory period appears to be necessary to complete the priming of dendritic spines that were stimulated, but not potentiated, by the first theta-burst stimulus. Priming allows these spines to potentiate after the second stimulus, and thus constitutes a biochemical stimulus trace (FIGS 1d,2a). Kramár *et al.*[47] noted that in hippocampal slices, additional potentiation can be induced up to 4 hours after induction of the first LTP increment[81]. The stimulus trace associated with priming may therefore take at

least 4 hours to decay. Such a long trace lifetime might allow a broad temporal window for optimal training trials.

In rat hippocampal slices, theta-burst stimuli lead to proteolytic inactivation of integrin receptors at stimulated dendritic spines[82]. These receptors are then replaced by vesicular transport of new receptors over a period of ~40–60 minutes, and it is hypothesized[82] that subsequent theta-burst stimuli at these synaptic contacts cannot induce spine enlargement or LTP until after this replacement has occurred, thus accounting for the refractory period of ~1 hour in order for a second theta-burst stimulus to yield additional LTP. This receptor replacement may constitute, at least in part, the priming of dendritic spines discussed above. These hypothesized dynamics may be in line with deficient-processing theory, with receptor replacement being the necessary process that can only occur during spaced inter-trial intervals (FIG. 3).

Transcription factor activation also constitutes a biochemical trace, and in some systems training may only be effective if inter-trial intervals are long enough so that each trial can induce a separate round of transcription and translation. Similarly, short (massed) inter-trial intervals may not lead to sufficient levels, or a sufficient duration, of activated MAPK or other kinases to support the consolidation of long-term memory.

The variant of deficient-processing theory positing that only spaced trials can generate sufficient cognitive rehearsals or reactivations of a memory to support long-term memory consolidation may also correspond to the empirical finding that repeated theta-burst stimuli, spaced by ~1 hour, can recruit additional dendritic spines by potentiating spines that were primed by preceding stimuli. A memory reactivation would be analogous to a theta-burst stimulus in that both events would initiate priming and potentiation. It also seems plausible that repeated memory reactivations might induce further rounds of transcription of genes involved in LTP, such as *C/ebp* and other CREB-activated genes, supporting further consolidation of long-term memory.

To more strongly connect this variant of deficient-processing theory to recent cellular and molecular data, one must also assume that reactivations of a memory reactivate some of the same neurons and synapses that were activated in the original learning sessions. In that way, the rehearsals and learning trials would reinforce memory in the same way. This assumption seems plausible but requires further empirical investigation. Although finer-grain analyses are necessary, a study using functional MRI during verbal learning supports this assumption[83]. In this study, a specific brain region associated with rehearsal of verbal memory, the left frontal operculum, was activated more during spaced learning of paired-word associations than during massed learning. We note that these posited memory reactivations, on timescales of ~1 hour or longer, are distinct from voluntary rehearsals of a memory on a short timescale (seconds or ~1 minute). Substantial behavioural evidence suggests that this latter voluntary, short-term rehearsal is not essential for spaced learning[31,34].

The remaining variants of the deficient-processing theory, which focus on habituation or on a lack of voluntary attention during massed presentations, do not appear to relate as readily

to the current single-neuron data. These variants have also been argued not to readily accommodate certain verbal learning observations[2]. With regards to encoding variability theory, data on neuronal network dynamics, rather than single-neuron data, will be needed to determine to what extent the binding of contexts to memory occurs, which is required in this theory. Similar data will also be needed to assess whether the binding of memories of later trials to those of earlier trials occurs, which is required in study-phase retrieval theory. It will be important to reassess all of these competing spaced learning theories as more information becomes available on the dynamics of memory networks. Indeed, different theories may be more or less applicable to different memory systems.

Irregular spacing can enhance learning

Attempts to optimize the spacing effect have generally been based on trial-and-error approaches. Consequently, most, if not all, training protocols used in animal and human studies are probably not optimal. For almost all learning paradigms, the training intervals are fixed, although in one type of spaced training paradigm, the intervals between sessions progressively lengthen[2,84]. However, a meta-analysis[2] and a text learning study[84] found no substantial evidence for the superiority of this approach in terms of promoting long-term memory formation.

It seems to be evident that at least part of the improvement in learning that is found with spaced training protocols can be explained by the dynamic relationships between the training trials and the underlying cellular and molecular mechanisms that are associated with memory formation (FIG. 3). But is the inverse possible? Can knowledge of the dynamics of the memory mechanisms be used to enhance memory processing by predicting optimal training protocols, possibly with irregular training intervals? One approach is to develop models of the biochemical cascades that underlie memory formation and use simulations to rapidly test the effectiveness of different training protocols[85]. In recent years, models have described the dynamics of the biochemical reactions that transduce stimuli into LTP[86–88]. These models have differential equations that simulate and predict the dynamics of the activities of key molecular species. Simulations have reproduced the dynamics of MAPK during LTP induction[80,86,88]. Models have also simulated the activity time courses of PKA, calcium/calmodulin-dependent protein kinase II (CaMKII), other key enzymes and downstream transcription factors during LTP induction[88–90]. Each signalling cascade in these models displays a characteristic activity time course; thus, it is likely some irregular sequence of intervals would be predicted to maximize the induction of LTP. For example, subsequent trials that are delivered at times that coincide with kinase activity peaks might optimally reinforce learning.

One study from our laboratory developed a model describing the 5-HT-induced PKA and ERK signalling pathways that are essential for LTF in *A. californica*[85]. In the model (FIG. 4a), the necessity of PKA and ERK activation for LTF was simply represented with a variable termed ‘inducer’. The value of inducer was proportional to the product of PKA and MAPK activities. The amount of LTF and LTS was predicted to increase with an increase in the peak value of inducer. Ten thousand different protocols consisting of five trials that were separated by intervals of 0–45 minutes were simulated (FIG. 4b). The ability to simulate and

predict the effects of so many protocols in a relatively short space of time represents a distinct advantage of computational studies over empirical studies. The simulations determined that, of these protocols, a massed protocol (FIG. 4b) produced the lowest peak value of inducer, consistent with data that massed 5-HT application fails to produce LTF[14]. The 'best' protocol, yielding the highest peak value for inducer, termed the 'enhanced' protocol (FIG. 4b), had irregular intervals. The protocol termed the 'standard' protocol (five 5-minute pulses of 5-HT, with uniform inter-stimulus intervals of 20 minutes) (FIG. 4b) has been commonly used to induce LTF in empirical studies for ~30 years[91]. This standard protocol yielded an intermediate peak value of inducer and was predicted to have an intermediate effectiveness (FIG. 4c). These predictions were empirically validated. The magnitude of LTF and LTS produced by the enhanced protocol exceeded that produced by the standard protocol[85] (FIG. 4d). An explanation for this enhancement of LTF, consistent with data[11,12,92], is as follows. In response to each 5-HT pulse, PKA activity increases rapidly and decays rapidly (FIG. 4c). MAPK activity rises and decays more slowly, not peaking until ~45 minutes after a pulse. The four initial pulses initiate a surge of MAPK activity, which peaks near the time of the last pulse. This last pulse activates PKA, so that a PKA peak is approximately coincident with peak MAPK activation, maximizing inducer and the predicted LTF.

If irregularly spaced protocols can enhance normal learning, might modelling also predict protocols capable of restoring learning that is impaired by a genetic mutation or other physiological insults? A recent study[90] tested this hypothesis. CREB-binding protein (CBP) is an acetyltransferase and essential co-activator for several transcription factors, including phosphorylated CREB (pCREB). CBP is also required for the consolidation of long-term memory[93]. Mutations that decrease CBP activity cause a human genetic disorder termed Rubinstein–Taybi syndrome (RTS)[94], which is associated with intellectual disability and learning deficits, and $Cbp^{+/-}$ mice show impaired LTP and long-term memory[95]. The recent study[90] used small-interfering RNA (siRNA) knockdown of CBP in *A. californica* sensory neurons to impair LTF. In this study, the model previously used[85] to predict optimal, irregularly spaced protocols that would enhance LTF (FIG. 4a) was extended to represent induction of *c/ebp*, a transcription factor known to be essential for LTF[96]. In simulations of the effects of different spaced protocols, greater peak levels of phosphorylated C/EBP (pC/EBP) were taken to predict greater LTF. Simulations showed a substantial decrease in pC/EBP levels when the level of CBP was reduced by a decrement that corresponded to the siRNA effect. A 'rescue' protocol with irregularly spaced intervals was predicted to restore peak pC/EBP levels and, correspondingly, LTF. This rescue protocol was empirically validated to restore normal LTF in *A. californica*. A similar predicted rescue protocol of irregularly spaced intervals rescued a deficit in LTF that was produced by siRNA knockdown of CREB1 (REF. 97).

Although these empirical studies were conducted in *A. californica*, it should be noted that key molecular mechanisms of memory are substantially conserved from simple model organisms such as *A. californica* to mammals[52,96]. For example, LTF and LTP both rely on PKA and ERK activation[12,92,98] and both rely on cooperative gene induction by pCREB and CBP[53,96,99,100]. LTF relies on deactivation of CREB2, a transcriptional repressor[101]. Similarly, relief of transcriptional repression owing to ATF4, a mammalian analogue of

CREB2, seems to be important for the maintenance of hippocampal LTP[102,103]. Thus, the results with *A. californica* suggest that it may be possible, in complex organisms including mammals, to computationally predict the efficacies of numerous learning or training protocols, a process that is impractical using empirical studies alone.

Given that knowledge of the underlying biochemical cascades can help to develop models to predict optimal training protocols, can models also be used to predict pharmacological targets to improve memory? The time may also be right for such an approach. For example, if simulated LTP deficits were rescued by combined parameter changes corresponding to known drug effects, these 'best' parameter combinations might prioritize drug combinations for testing in animal models. A recent study[104] took a first step by modelling LTP induction and transcriptional regulation by CREB, and simulating the effects of drugs on LTP by altering the parameters of the model. In this model, the magnitude of LTP induction was represented by an increase in a synaptic weight variable. LTP impairment seen in a mouse model of RTS[95] was first simulated. Then, starting from this simulation, the parameters were altered in ways corresponding to plausible single-drug effects. However, no single-drug effect completely rescued LTP. Thus, pairs of parameter changes were considered, corresponding to plausible paired-drug effects. Two pairs were identified that restored LTP. In the first case, an increased rate constant for histone acetylation, corresponding to application of an acetyltransferase activator, was paired with an increased duration of stimulus-induced increase in cAMP levels, corresponding to application of a cAMP phosphodiesterase (PDE) inhibitor. The second pair corresponded to a PDE inhibitor paired with a deacetylase inhibitor. For both pairs, additive drug synergism, defined as a combined-drug effect that exceeds the summed effects of the individual drugs, was also evident, as quantified by a simple additive measure (FIG. 5). A subsequent empirical study by another group did find that pairing a PDE inhibitor with a deacetylase inhibitor was effective in rescuing a deficit of LTP in a mouse model of Alzheimer disease[105]. A further extension of these strategies might similarly predict, and empirically test, enhancement of synaptic plasticity when pharmacotherapy is combined with computationally designed spaced protocols.

Future directions

There is reason for optimism that more predictive models for determining optimal intervals between learning trials will be available in the near future, because the molecular data that are necessary for the development of such models, which can delineate the dynamics of signalling pathways that are important for LTP and long-term memory, continue to accumulate rapidly. However, despite the progress being made in understanding the molecular mechanisms of the spacing effect, some aspects of this effect cannot be explained by current models and constitute important directions for future research. For example, in human verbal learning, an interesting positive correlation exists between the length of inter-trial intervals for effective spaced learning and the retention interval (that is, the interval between the final training trial and the test of memory retention). With relatively short retention intervals (~1 minute−2 hours), training intervals in the broad range of ~1 minute to 3 hours yield greater verbal learning than do training intervals of 2 days or more[2]. With a longer retention interval of 1 day, a 1-day training interval yielded greater learning than did a

very short (<30-second) interval. For verbal learning with a retention interval of 6 months, a training interval of 7 days was superior to an interval of 3 days[71]. This correlation between longer training and retention intervals suggests that longer training intervals preferentially form a memory trace with a very long lifetime. For the temporal range of minutes versus hours, it is plausible that a longer trace lifetime corresponds, at least in part, to increased activation of transcription by the longer training intervals. However, this explanation may not suffice when comparing training intervals of ~1 day versus many days. It would be of interest to determine whether reactivation of stored memory representations at the network level, or transfer of these representations between brain regions, contributes to this correlation.

Another challenge will be to use innovative strategies to test the predictions of the cognitive theories for the spacing effect. For example, consider the variant of deficient-processing theory positing that repeated cognitive rehearsals of a memory are needed for consolidation. A neuronal correlate of rehearsals is, plausibly, repeated activation of a specific neuron assembly that serves as a locus of storage of a long-term memory trace. Empirically, is such repeated activation necessary for persistence of memory for days or longer? Repeated spontaneous activation of neuron assemblies does occur[106,107], as does repeated replay or rehearsal of assemblies that encode recent experiences[108,109]. One study supporting the necessity of such replay found that the post-training suppression of activity of neurons that were engineered to overexpress CREB in the amygdala blocked the consolidation of a memory of association between cocaine and a location[110]. Similar blocking effects were obtained by the indiscriminate activation of neurons that overexpressed CREB. Although encouraging, these manipulations lack the cellular precision that is necessary to demonstrate conclusively that reactivation of a particular assembly of neurons is essential for the persistence of long-term memory. Future studies using optogenetic techniques could provide that precision. Similarly, innovative strategies will be needed to address whether effective spaced learning requires the binding of contextual and episodic memories at the neuronal network level, such as posited by encoding variability theory, or increased binding due to greater retrieval effort, as posited by study-phase retrieval theory.

The successful prediction of the interval structure of behavioural training protocols that may overcome some human learning deficits (when applied alone or in combination with pharmacotherapy) will require improved knowledge of the signalling pathways that underlie LTP and long-term memory formation and of the ways in which the deficits affect those pathways. Future models are still likely to be incomplete owing to gaps in knowledge. For example, data will be incomplete and associated with unavoidable uncertainties in the values of biochemical parameters such as enzyme activities or protein concentrations. In model development, data from several preparation types (for example, cell cultures and slices) and species (for example, primates and rodents) commonly need to be used to estimate different parameters[86,88]. However, although these limitations are important, the potential benefits of combining modelling with experiments in the ways discussed in this Review are extensive, such that this strategy may have promise for improving the clinical and educational outcomes for patients with learning and memory deficits. In addition, it is possible that education and learning in individuals without such deficits could benefit from such a strategy. Indeed, enhancing normal learning by judicious pharmacotherapy has recently

received attention[111], and combining drugs with optimized spaced learning protocols might yield even better outcomes.

Acknowledgements

This work was supported by US National Institutes of Health grants NS073974 and NS019895.

Glossary

Memory extinction	The decline of a learned behavioural response to a conditioned stimulus following the withdrawal of reinforcement stimuli that were previously paired with repetitions of the conditioned stimulus
Reinforcement	A broad term used here to describe a stimulus or item that enhances the strength or lifetime of a memory
Habituation	A decrease in the behavioural response to a stimulus following frequent repetitions of that stimulus; this term is distinct from extinction, because habituation can denote a decrease in response to a stimulus that was never paired with a reinforcing stimulus
Memory reactivations	These are reinstatements of conditioned behavioural responses or of neural activity associated with specific responses and can be elicited by presentation of a conditioning stimulus or of the context in which learning previously occurred, or be spontaneous, occurring as a part of normal ongoing neural activity
Drug synergism	In combined-drug treatment, a synergistic effect of the combination is an effect that is greater than that which would be predicted by considering the individual drugs as independent and not interacting

References

1. Ebbinghaus, H. Memory; a Contribution to Experimental Psychology. Teachers College, Columbia University; 1913. Ch. 2
2. Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: a review and quantitative synthesis. Psychol. Bull. 2006; 132:354–380. This extensive review and meta-analysis delineates the comprehensive body of knowledge describing human spaced and massed verbal learning, as well as theories posited to explain the superiority of spaced training over massed training in terms of inducing long-term memory formation. [PubMed: 16719566]
3. Godbole NR, Delaney PF, Verkoeijen PP. The spacing effect in immediate and delayed free recall. Memory. 2014; 22:462–469. [PubMed: 23701389]
4. Raman M, et al. Teaching in small portions dispersed over time enhances long-term knowledge retention. Med. Teach. 2010; 32:250–255. [PubMed: 20218841]
5. Donovan JJ, Radosevich DJ. A meta-analytic review of the distribution of practice effect. J. Appl. Psychol. 1999; 84:795–805.

6. Shea CH, Lai Q, Black C, Park JH. Spacing practice sessions across days benefits the learning of motor skills. Hum. Mov. Sci. 2000; 19:737–760.
7. Gluckman M, Vlach HA, Sandhofer CM. Spacing simultaneously promotes multiple forms of learning in children's science curriculum. Appl. Cognit. Psychol. 2014; 28:266–273.
8. McDaniel, MA.; Callender, AA. Learning and Memory: A Comprehensive Reference. Roediger, HL., editor. Vol. 2. Academic Press; 2008. p. 819-844.
9. Sobel HS, Cepeda NJ, Kapler IV. Spacing effects in real-world classroom vocabulary learning. Appl. Cognit. Psychol. 2011; 25:763–767.
10. Vlach HA. The spacing effect in children's generalization of knowledge: allowing children time to forget promotes their ability to learn. Child Dev. Persp. 2014; 8:163–168. This study found that, in children's education, spacing of learning sessions promoted not only better fact retention but also generalization of currently learned items to bind with and extend previously established conceptual categories.
11. Philips GT, Kopec AM, Carew TJ. Pattern and predictability in memory formation: from molecular mechanisms to clinical relevance. Neurobiol. Learn. Mem. 2013; 105:117–124. [PubMed: 23727358]
12. Philips GT, Ye X, Kopec AM, Carew TJ. MAPK establishes a molecular context that defines effective training patterns for long-term memory formation. J. Neurosci. 2013; 33:7565–7573. This work characterized the dynamics of MAPK and its activation of C/EBP, a transcription factor that is important for long-term memory formation, identifying the response of MAPK within a specific time window as essential for generating an optimal inter-trial interval. [PubMed: 23616561]
13. Wainwright ML, Zhang H, Byrne JH, Cleary LJ. Localized neuronal outgrowth induced by long-term sensitization training in *Aplysia*. J. Neurosci. 2002; 22:4132–4141. [PubMed: 12019331]
14. Mauelshagen J, Sherff CM, Carew TJ. Differential induction of long-term synaptic facilitation by spaced and massed applications of serotonin at sensory neuron synapses of *Aplysia californica*. Learn. Mem. 1998; 5:246–256. [PubMed: 10454368]
15. Beck CD, Schroeder B, Davis RL. Learning performance of normal and mutant *Drosophila* after repeated conditioning trials with discrete stimuli. J. Neurosci. 2000; 20:2944–2953. [PubMed: 10751447]
16. Pagani MR, Oishi K, Gelb BD, Zhong Y. The phosphatase SHP2 regulates the spacing effect for long-term memory induction. Cell. 2009; 139:186–198. [PubMed: 19804763]
17. Menzel R, Manz G, Menzel R, Greggers U. Massed and spaced learning in honeybees: the role of CS, US, the intertrial interval, and the test interval. Learn. Mem. 2001; 8:198–208. [PubMed: 11533223]
18. Anderson MJ, Jablonski SA, Klimas DB. Spaced initial stimulus familiarization enhances novelty preference in Long-Evans rats. Behav. Processes. 2008; 78:481–486. [PubMed: 18358637]
19. Bello-Medina PC, Sanchez-Carrasco L, Gonzalez-Ornelas NR, Jeffery KJ, Ramirez-Amaya V. Differential effects of spaced versus massed training in long-term object-identity and object-location recognition memory. Behav. Brain Res. 2013; 250:102–113. [PubMed: 23644160]
20. Medin DL. The comparative study of memory. J. Hum. Evol. 1974; 3:455–463.
21. Robbins O, Bush CT. Memory in great apes. J. Exp. Psychol. 1973; 97:344–348.
22. Urcelay GP, Wheeler DS, Miller RR. Spacing extinction trials alleviates renewal and spontaneous recovery. Learn. Behav. 2009; 37:60–73. [PubMed: 19122053]
23. Benjamin AS, Tullis J. What makes distributed practice effective? Cogn. Psychol. 2010; 61:228–247. [PubMed: 20580350]
24. Madigan SA. Intraserial repetition and coding processes in free recall. J. Verb. Learn. Verb. Behav. 1969; 8:828–835.
25. Melton AW. The situation with respect to the spacing of repetitions and memory. J. Verb. Learn. Verb. Behav. 1970; 9:596–606.
26. Braun K, Rubin DC. The spacing effect depends on an encoding deficit, retrieval, and time in working memory: evidence from once-presented words. Memory. 1998; 6:37–65. [PubMed: 9640432]

27. Greene RL. Spacing effects in memory: evidence for a two-process account. J. Exp. Psychol. Learn. Mem. Cogn. 1989; 15:371–377.
28. Hintzman DL, Summers JJ, Block RA. Spacing judgments as an index of study-phase retrieval. J. Exp. Psychol. Hum. Learn. 1975; 1:31–40.
29. Tzeng OJ, Cotton B. A study-phase retrieval model of temporal coding. J. Exp. Psychol. Hum. Learn. 1980; 6:705–716.
30. Siegel LL, Kahana MJ. A retrieved context account of spacing and repetition effects in free recall. J. Exp. Psychol. Learn. Mem. Cogn. 2014; 40:755–764. [PubMed: 24564545]
31. Hintzman, DL. Theories in Cognitive Psychology: The Loyola Symposium. Solso, RL., editor. Erlbaum; 1974. p. 77-90.
32. Dempster FN. Spacing effects and their implications for theory and practice. Educ. Psychol. Rev. 1989; 1:309–330.
33. Toppino TC. The spacing effect in young children's free recall: support for automatic-process explanations. Mem. Cognit. 1991; 19:159–167.
34. Hintzman, DL.; Bower, GH. The Psychology of Learning and Motivation. Vol. 10. Academic Press; 1976. p. 47-91.
35. Rundus D. Analysis of rehearsal processes in free recall. J. Exp. Psychol. 1971; 89:63–77.
36. Lechner HA, Squire LR, Byrne JH. 100 years of consolidation — remembering Müller and Pilzecker. Learn. Mem. 1999; 6:77–87. [PubMed: 10327233]
37. Landauer TK. Reinforcement as consolidation. Psychol. Rev. 1969; 76:82–96. This early paper posited and described elements of the dynamics of spaced learning, such as the saturation of memory formation by closely spaced trials and the failure of reinforcement for long trial spacing, and how these elements generate an optimal inter-trial interval. [PubMed: 5353380]
38. Peterson LR. Short-term verbal memory and learning. Psychol. Rev. 1966; 73:193–207. [PubMed: 5325895]
39. Wickelgren WA. Trace resistance and the decay of long-term memory. J. Math. Psychol. 1972; 9:418–455.
40. Cuddy LJ, Jacoby LJ. When forgetting helps memory: an analysis of repetition effects. J. Verb. Learn. Verb. Behav. 1982; 21:451–467.
41. Toppino TC, Bloom LC. The spacing effect, free recall, and two-process theory: a closer look. J. Exp. Psychol. Learn. Mem. Cogn. 2002; 28:437–444. [PubMed: 12018496]
42. Nabavi S, et al. Engineering a memory with LTD and LTP. Nature. 2014; 511:348–352. [PubMed: 24896183]
43. Bosch M, Hayashi Y. Structural plasticity of dendritic spines. Curr. Opin. Neurobiol. 2012; 22:383–388. [PubMed: 21963169]
44. Bell ME, et al. Dynamics of nascent and active zone ultrastructure as synapses enlarge during LTP in mature hippocampus. J. Comp. Neurol. 2014; 522:3861–3864. This work, one of a group of related studies by the Harris laboratory, characterized in detail pre- and postsynaptic structural changes that are necessary for the consolidation of LTP. [PubMed: 25043676]
45. Bourne JN, Harris KM. Coordination of size and number of excitatory and inhibitory synapses results in a balanced structural plasticity along mature hippocampal CA1 dendrites during LTP. Hippocampus. 2011; 21:354–373. [PubMed: 20101601]
46. Cao G, Harris KM. Augmenting saturated LTP by broadly spaced episodes of theta-burst stimulation in hippocampal area CA1 of adult rats and mice. J. Neurophysiol. 2014; 112:1916–1924. [PubMed: 25057146]
47. Kramár EA, et al. Synaptic evidence for the efficacy of spaced learning. Proc. Natl Acad. Sci. USA. 2012; 109:5121–5126. This study, together with reference 48, substantiates the key hypothesis that spaced stimuli first prime and then strengthen upon repetition successive groups of dendritic spines, creating a morphological correlate of the memory trace that spaced, but not massed, protocols effectively consolidate. [PubMed: 22411798]
48. Lynch G, Kramár EA, Babayan AH, Rumbaugh G, Gall CM. Differences between synaptic plasticity thresholds result in new timing rules for maximizing long-term potentiation. Neuropharmacology. 2013; 64:27–38. [PubMed: 22820276]

49. Rex CS, et al. Different Rho GTPase-dependent signaling pathways initiate sequential steps in the consolidation of long-term potentiation. J. Cell Biol. 2009; 186:85–97. [PubMed: 19596849]
50. Nguyen PV, Abel T, Kandel ER. Requirement of a critical period of transcription for induction of a late phase of LTP. Science. 1994; 265:1104–1107. [PubMed: 8066450]
51. Scharf MT, et al. Protein synthesis is required for the enhancement of long-term potentiation and long-term memory by spaced training. J. Neurophysiol. 2002; 87:2770–2777. [PubMed: 12037179]
52. Kandel ER. The molecular biology of memory storage: a dialogue between genes and synapses. Science. 2001; 294:1249–1254.
53. Silva AJ, Kogan JH, Frankland PW, Kida S. CREB and memory. Annu. Rev. Neurosci. 1998; 21:127–148. [PubMed: 9530494]
54. Liu RY, Fioravante D, Shah S, Byrne JH. cAMP response element-binding protein 1 feedback loop is necessary for consolidation of long-term synaptic facilitation in *Aplysia*. J. Neurosci. 2008; 28:1970–1976. [PubMed: 18287513]
55. Liu RY, Cleary LJ, Byrne JH. The requirement for enhanced CREB1 expression in consolidation of long-term synaptic facilitation and long-term excitability in sensory neurons of *Aplysia*. J. Neurosci. 2011; 31:6871–6879. [PubMed: 21543617]
56. Liu RY, Shah S, Cleary LJ, Byrne JH. Serotonin- and training-induced dynamic regulation of CREB2 in *Aplysia*. Learn. Mem. 2011; 18:245–249. [PubMed: 21441301]
57. Frost WN, Castellucci V, Hawkins RD, Kandel ER. Monosynaptic connections made by the sensory neurons of the gill- and siphon-withdrawal reflex in *Aplysia* participate in the storage of long-term memory for sensitization. Proc. Natl Acad. Sci. USA. 1985; 82:8266–8269. [PubMed: 16593630]
58. Pinsker HM, Hening WA, Carew TJ, Kandel ER. Long-term sensitization of a defensive withdrawal reflex in *Aplysia*. Science. 1973; 182:1039–1042. [PubMed: 4748675]
59. Beckinschtein P, et al. Persistence of long-term memory storage requires a late protein synthesis- and BDNF-dependent phase in the hippocampus. Neuron. 2007; 53:261–277. [PubMed: 17224407]
60. Bambah-Mukku D, Travaglia A, Chen DY, Pollonini G, Alberini CM. A positive autoregulatory BDNF feedback loop via C/EBPβ mediates hippocampal memory consolidation. J. Neurosci. 2014; 34:12547–12559. This study characterizes the dynamics of a molecular positive feedback loop demonstrated to be important for the consolidation of long-term (24 hours or more) memory, as well as the late repression of *Bdnf* that terminates the feedback. [PubMed: 25209292]
61. Genoux D, et al. Protein phosphatase 1 is a molecular constraint on learning and memory. Nature. 2002; 418:970–974. [PubMed: 12198546]
62. Farah CA, Weatherill D, Dunn TW, Sossin WS. PKC differentially translocates during spaced and massed training in *Aplysia*. J. Neurosci. 2009; 29:10281–10286. [PubMed: 19692602]
63. Tully T, Preat T, Boynton S, Del Vecchio M. Genetic dissection of consolidated memory in *Drosophila melanogaster*. Cell. 1994; 79:35–47. [PubMed: 7923375]
64. Yin JC, Del Vecchio M, Zhou H, Tully T. CREB as a memory modulator: induced expression of a dCREB2 activator isoform enhances long-term memory in *Drosophila*. Cell. 1995; 81:107–115. [PubMed: 7720066]
65. Smolen P, Baxter DA, Byrne JH. Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems. Am. J. Physiol. 1998; 274:C531–C542. [PubMed: 9486144]
66. Philips GT, Tzvetkova EI, Carew TJ. Transient mitogen-activated protein kinase activation is confined to a narrow temporal window required for the induction of two-trial long term memory in *Aplysia*. J. Neurosci. 2007; 27:13701–13705. [PubMed: 18077681]
67. Seese RR, Wang K, Yao YQ, Lynch G, Gall CM. Spaced training rescues memory and ERK1/2 signaling in fragile X syndrome model mice. Proc. Natl Acad. Sci. USA. 2014; 111:16907–16912. [PubMed: 25385607]
68. Rosenblum K, et al. The role of extracellular regulated kinases I/II in late-phase long-term potentiation. J. Neurosci. 2002; 22:5432–5441. [PubMed: 12097495]
69. Ajay SM, Bhalla US. A propagating ERKII switch forms zones of elevated dendritic activation correlated with plasticity. HFSP J. 2007; 1:49–66. [PubMed: 19404460]

70. Chen C, et al. Visualizing long-term memory formation in two neurons of the *Drosophila* brain. Science. 2012; 335:678–685. [PubMed: 22323813]
71. Cepeda NJ, et al. Optimizing distributed practice: theoretical analysis and practical implications. Exp. Psychol. 2009; 56:236–246. [PubMed: 19439395]
72. Wei Y, Koulakov AA. Long-term memory stabilized by noise-induced rehearsal. J. Neurosci. 2014; 34:15804–15815. [PubMed: 25411507]
73. Cui Z, et al. Inducible and reversible NR1 knockout reveals a crucial role of the NMDA receptor in preserving remote memories in the brain. Neuron. 2004; 41:781–793. This study demonstrates the key point that ongoing neural activity and NMDA receptor activation are essential for long-lasting preservation of stored memory. [PubMed: 15003177]
74. Tetzlaff C, Kolodziejski C, Timme M, Tsodyks M, Worgotter F. Synaptic scaling enables dynamically distinct short- and long-term memory formation. PLoS Comput. Biol. 2013; 9:e1003307. [PubMed: 24204240]
75. Miller P, Zhabotinsky AM, Lisman JE, Wang XJ. The stability of a stochastic CaMKII switch: dependence on the number of enzyme molecules and protein turnover. PLoS Biol. 2005; 3:e107. [PubMed: 15819604]
76. Pettigrew DB, Smolen P, Baxter DA, Byrne JH. Dynamic properties of regulatory motifs associated with induction of three temporal domains of memory in *Aplysia*. J. Comput. Neurosci. 2005; 18:163–181. [PubMed: 15714268]
77. Shema R, Sacktor TC, Dudai Y. Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKMζ. Science. 2007; 317:951–953. [PubMed: 17702943]
78. Volk LJ, Bachman JL, Johnson R, Yu Y, Huganir RL. PKM-ζ is not required for hippocampal synaptic plasticity, learning and memory. Nature. 2013; 493:420–423. [PubMed: 23283174]
79. Fioriti L, et al. The persistence of hippocampal-based memory requires protein synthesis mediated by the prion-like protein CPEB3. Neuron. 2015; 86:1433–1448. [PubMed: 26074003]
80. Ajay SM, Bhalla US. A role for ERKII in synaptic pattern selectivity on the time scale of minutes. Eur. J. Neurosci. 2004; 20:2671–2680. [PubMed: 15548210]
81. Frey U, Morris RG. Synaptic tagging: implications for late maintenance of hippocampal long-term potentiation. Trends Neurosci. 1998; 21:181–188. [PubMed: 9610879]
82. Babayan AH, et al. Integrin dynamics produce a delayed stage of long-term potentiation and memory consolidation. J. Neurosci. 2012; 32:12854–12861. [PubMed: 22973009]
83. Callan DE, Schweighofer N. Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. Hum. Brain Mapp. 2010; 31:645–659. [PubMed: 19882649]
84. Karpicke JD, Roediger HL. Is expanding retrieval a superior method for learning text materials? Mem. Cognit. 2010; 38:116–124.
85. Zhang Y, et al. Computational design of enhanced learning protocols. Nat. Neurosci. 2012; 15:294–297. This study found that computational models describing long-term synaptic strengthening can successfully predict optimal spaced training protocols, with irregular inter-trial intervals, that enhance learning.
86. Bhalla US, Iyengar R. Emergent properties of networks of biological signaling pathways. Science. 1999; 283:381–387. [PubMed: 9888852]
87. Kim M, Huang T, Abel T, Blackwell KT. Temporal sensitivity of protein kinase A activation in late-phase long term potentiation. PLoS Comput. Biol. 2010; 6:e1000691. [PubMed: 20195498]
88. Smolen P, Baxter DA, Byrne JH. A model of the roles of essential kinases in the induction and expression of late long-term potentiation. Biophys. J. 2006; 90:2760–2775. [PubMed: 16415049]
89. Aslam N, Shouval HZ. Regulation of cytoplasmic polyadenylation can generate a bistable switch. BMC Syst. Biol. 2012; 6:12–20. [PubMed: 22335938]
90. Liu RY, et al. Deficit in long-term synaptic plasticity is rescued by a computationally predicted stimulus protocol. J. Neurosci. 2013; 33:6944–6949. [PubMed: 23595752]
91. Montarolo PG, et al. A critical period for macromolecular synthesis in long-term heterosynaptic facilitation in *Aplysia*. Science. 1986; 234:1249–1254. [PubMed: 3775383]

92. Müller U, Carew TJ. Serotonin induces temporally and mechanistically distinct phases of persistent PKA activity in *Aplysia* sensory neurons. Neuron. 1998; 21:1423–1434. [PubMed: 9883734]

93. Korzus E, Rosenfeld MG, Mayford M. CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron. 2004; 42:961–972. [PubMed: 15207240]

94. Park E, et al. Epigenetic mechanisms of Rubinstein–Taybi syndrome. Neuromolecular Med. 2014; 16:16–24. [PubMed: 24381114]

95. Alarcon JM, et al. Chromatin acetylation, memory, and LTP are impaired in $CBP^{+/-}$ mice: a model for the cognitive deficit in Rubinstein–Taybi syndrome and its amelioration. Neuron. 2004; 42:947–959. [PubMed: 15207239]

96. Alberini CM. Transcription factors in long-term memory and synaptic plasticity. Physiol. Rev. 2009; 89:121–145. [PubMed: 19126756]

97. Zhou L, et al. Rescue of impaired long-term facilitation at sensorimotor synapses of *Aplysia* following siRNA knockdown of CREB1. J. Neurosci. 2015; 35:1617–1626. [PubMed: 25632137]

98. Abel T, Nguyen PV. Regulation of hippocampus-dependent memory by cyclic AMP-dependent protein kinase. Prog. Brain Res. 2008; 169:97–115. [PubMed: 18394470]

99. Bartsch D, Casadio A, Karl KA, Serodio P, Kandel ER. CREB1 encodes a nuclear activator, a repressor, and a cytoplasmic modulator that form a regulatory unit critical for long-term facilitation. Cell. 1998; 95:211–223. [PubMed: 9790528]

100. Kida S. A functional role for CREB as a positive regulator of memory formation and LTP. Exp. Neurobiol. 2012; 21:136–140. [PubMed: 23319873]

101. Bartsch D, et al. *Aplysia* CREB2 represses long-term facilitation: relief of repression converts transient facilitation into long-term functional and structural change. Cell. 1995; 83:979–992. [PubMed: 8521521]

102. Dong C, Upadhya SC, Ding L, Smith TK, Hegde AN. Proteosome inhibition enhances the induction and impairs the maintenance of late-phase long-term potentiation. Learn. Mem. 2008; 15:335–347. [PubMed: 18441292]

103. Pasini S, Corona C, Liu J, Greene LA, Shelanski ML. Specific downregulation of hippocampal ATF4 reveals a necessary role in synaptic plasticity and memory. Cell Rep. 2015; 11:183–191. [PubMed: 25865882]

104. Smolen P, Baxter DA, Byrne JH. Simulations suggest pharmacological methods for rescuing long-term potentiation. J. Theor. Biol. 2014; 360:243–250. [PubMed: 25034337]

105. Cuadrado-Tejedor M, et al. Concomitant histone deacetylase and phosphodiesterase 5 inhibition synergistically prevents the disruption in synaptic plasticity and it reverses cognitive impairment in a mouse model of Alzheimer's disease. Clin. Epigenetics. 2015; 7:108. [PubMed: 26457123]

106. Ikegaya Y, et al. Synfire chains and cortical songs: temporal modules of cortical activity. Science. 2004; 304:559–564. [PubMed: 15105494]

107. Miller JE, Ayzenshtat I, Carrillo-Reid L, Yuste R. Visual stimuli recruit intrinsically generated cortical ensembles. Proc. Natl Acad. Sci. USA. 2014; 111:E4053–E4061. [PubMed: 25201983]

108. Karlsson MP, Frank LM. Awake replay of remote experiences in the hippocampus. Nat. Neurosci. 2009; 12:913–919. [PubMed: 19525943]

109. Wu X, Foster DJ. Hippocampal replay captures the unique topological structure of a novel environment. J. Neurosci. 2014; 34:6459–6469. [PubMed: 24806672]

110. Hsiang HL, et al. Manipulating a "cocaine engram" in mice. J. Neurosci. 2014; 34:14115–14127. [PubMed: 25319707]

111. Lynch G, Cox CD, Gall CM. Pharmacological enhancement of memory or cognition in normal subjects. Front. Syst. Neurosci. 2014; 8:90. [PubMed: 24904313]

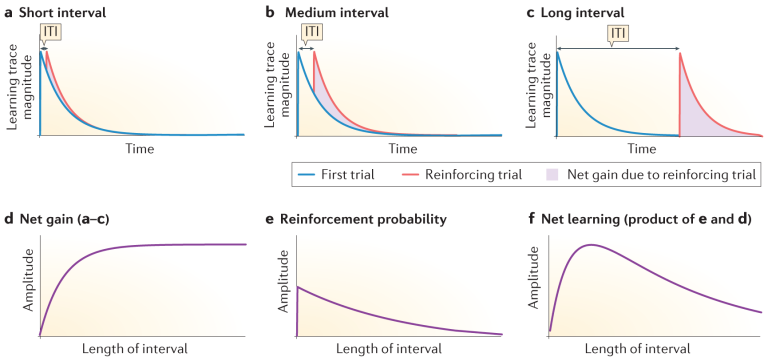

Figure 1. Early conceptual model of how learning trace dynamics generate an optimal interval

As described by the early model of Landauer[37], spaced training is more effective than massed training at strengthening some form of trace corresponding to memory storage in the brain, although this conceptual model does not posit a biochemical or structural form for the trace. This model posits that memory formation becomes more effective with longer inter-stimulus intervals between training sessions because of decreasing temporal overlap between successive, short-lived learning traces. These learning traces do not themselves constitute a memory. However, their net effect contributes to the formation of a long-lived memory trace. **a**–**c** | Learning traces elicited by two successive trials are shown. The model assumes that, for each value of the inter-trial interval (ITI) length, a quantity denoted ‘net gain owing to the reinforcing trial’ is proportional to the red area. Shorter intervals are associated with more overlap of learning traces and less net gain. Thus, a reinforcing trial is most effective after a refractory period following the preceding trial. For this conceptual model, units for amplitude and time are arbitrary. **d** | A greater summed effect, or net gain, of reinforcing trials occurs for longer inter-stimulus intervals. The effect reaches a plateau for long intervals as the overlap between successive learning traces reaches zero. **e** | Over longer times, a different quantity — the probability that a reinforcing trial will be effective at all in reactivating processes that constituted the preceding learning trace — declines. **f** | An optimum interval for maximizing the strength of the long-lived memory trace results when the greater net gain of reinforcement at longer intervals (from part **d**) is multiplied by the slowly declining probability that a reinforcement will reactivate a previous learning trace (from part **e**). The optimum interval for net learning is the one that produces the peak level of the trace in part **f**.

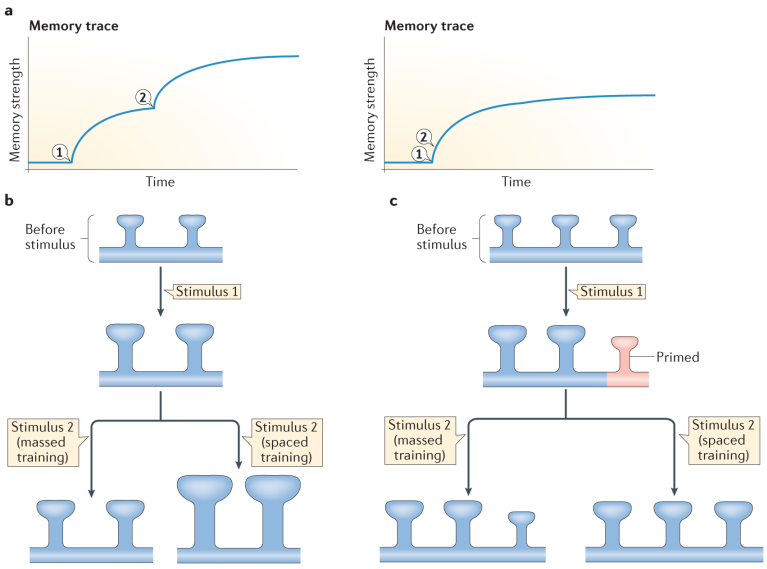

Figure 2. Model and hypotheses describing synaptic strengthening during spaced learning

a | In the refractory-state model, spaced stimuli (left panel; stimulus 1, followed substantially later by stimulus 2) cumulatively strengthen a memory trace (blue time course). By contrast, massed stimuli (right panel; stimulus 1 followed shortly after by stimulus 2) fail to cumulatively strengthen the memory trace. **b** | The cumulative synaptic strengthening in spaced training may be due to progressive enhancement of long-term potentiation (LTP), which could result from successive increases in the strength of the same synaptic contacts (shown here as successive increases in the volume of the same postsynaptic dendritic spine). Thus, in one of two current hypotheses describing synaptic strengthening during spaced learning, stimulus 1 enlarges a population of spines. If stimulus 2 follows shortly after the first stimulus (as in massed training), it cannot further affect spines. However, if stimulus 2 comes after a refractory period (as in spaced training), it can further enlarge the same population of spines. **c** | Alternatively, enhancement of LTP could result from successive rounds of strengthening of new synaptic contacts. Thus, in the second current hypothesis, stimulus 1 only enlarges a subset of affected spines, but primes additional spines. If stimulus 2 follows shortly after stimulus 1 (as in massed training), it has no effect. If stimulus 2 comes later (as in spaced training), it does not further enlarge the first subset of spines. Instead, stimulus 2 enlarges those spines that were primed, but not enlarged, by stimulus 1.

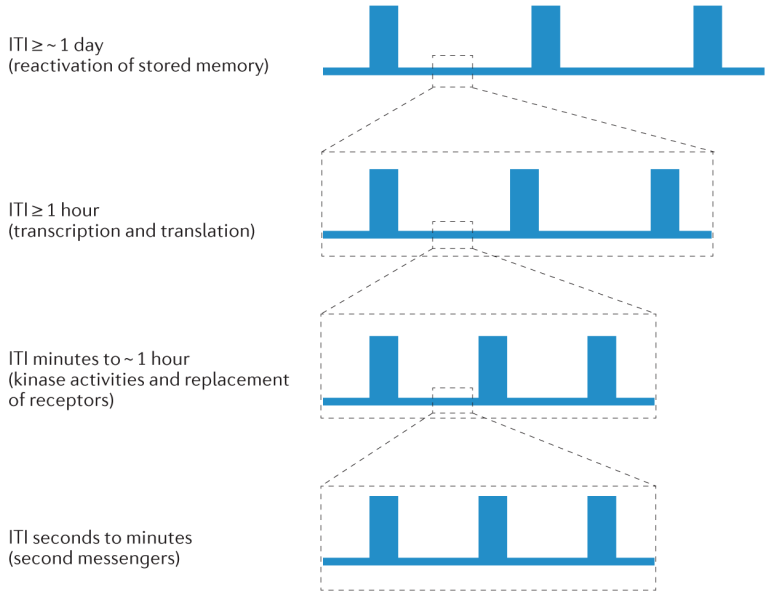

Figure 3. Different mechanisms may underlie enhancement of learning by spaced intervals of widely varying lengths

For relatively brief inter-trial intervals (ITIs) (bottom trace), successive trials may coincide with and reinforce peak second messenger levels generated by preceding trials. In each trace, individual rectangles represent individual trials, and converging lines between traces represent the lengthening of timescales as one moves upwards in the illustration. For somewhat longer ITIs (several minutes to ~1 hour), successive trials may reinforce the peak activities of kinases elicited by preceding trials and also elicit long-term potentiation of primed dendritic spines. Intervals of this length may also, in the hippocampus, be needed to allow replacement of inactivated receptors at stimulated spines[82], enabling succeeding stimulus repetitions to potentiate those spines. For intervals of ~1 hour or more, succeeding trials may also align with peaks in transcription factor activity and gene expression owing to preceding trials. For the longest ITIs (many hours or longer), succeeding trials may reactivate and thereby further potentiate consolidated memory traces. All of these processes are likely to contribute to the consolidation of long-term memory, in many if not all species. However, depending on the ITI length used in a particular spaced learning protocol, the dynamics of a particular type of process (for example, kinase activation) may contribute in particular to the efficacy of spaced learning. Also, trials at one temporal domain (for example, 1 day) may be unitary events, but also may constitute a block of spaced trials from another temporal domain (for example, minutes to hours). For example, an effective protocol for long-term sensitization training in *Aplysia californica* is the use of four trials with an ITI of 30 minutes, with this block repeated four times with a 1-day ITI[13]. Thus, some effective training protocols consist of a hierarchy of temporal domains of training sessions, with briefer sessions embedded within longer ones. In this illustration, intervals are shown with regular spacing, but more effective learning may occur with irregular spacing (FIG. 4).

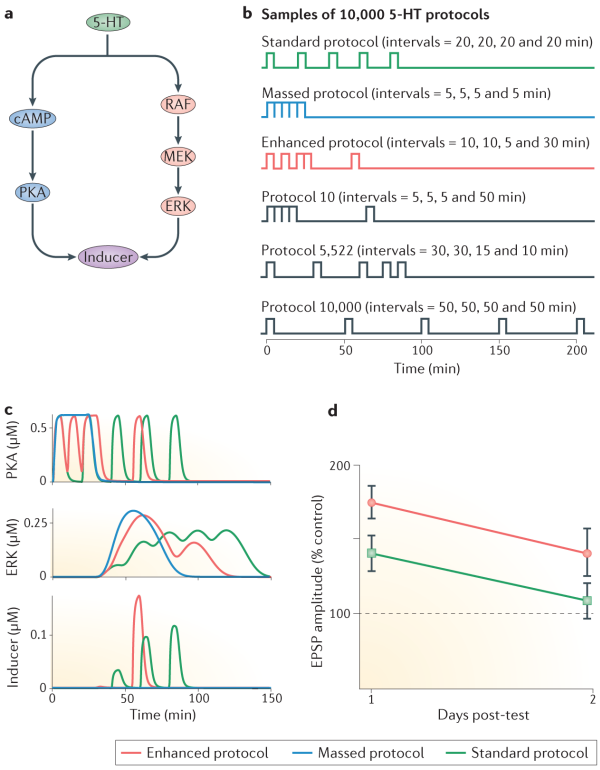

Figure 4. Dynamics of a model that has successfully predicted greater efficacy for a learning protocol with irregularly spaced intervals

a | A simplified mathematical model[85] describes the activation and effects of two key kinases necessary for long-term facilitation (LTF), a cellular correlate of a simple form of learning, long-term sensitization. Brief applications of 5-hydroxytryptamine (5-HT) activate protein kinase A (PKA) by increasing the levels of the secondary messenger cyclic AMP, and activate the extracellular signal-regulated kinase (ERK) isoform of mitogen-activated protein kinase (MAPK) via a RAS–RAF–MEK cascade. PKA and ERK interact, at least in part, via the phosphorylation of transcription factors, to induce LTF. In the model, the variable 'inducer' represents the PKA–ERK interaction. A higher peak value of inducer was assumed to predict a greater amplitude of LTF. **b** | Six samples of the 10,000 5-HT protocols that were simulated with the model. All protocols consist of five 5-minute pulses of 5-HT, shown as rectangular waves, with inter-pulse intervals chosen as multiples of 5 minutes, in the range of 5–50 minutes. The standard protocol (green trace) is the protocol most commonly used in studies of LTF *in vitro*. The enhanced protocol (red trace) produced the largest peak value of inducer, whereas the massed protocol (blue trace) produced the smallest peak value of inducer. The standard protocol has uniform inter-pulse intervals of 20 minutes, whereas the enhanced protocol has non-uniform intervals of 10, 10, 5 and 30 min. The massed protocol has no gaps between the 5-HT pulses. **c** | Simulated time courses of activated PKA, activated ERK and inducer in response to the standard protocol (green traces), the enhanced protocol (red traces) and the massed protocol (blue traces). **d** | In an empirical validation of the model's prediction, the LTF induced by the enhanced protocol, as determined by the percentage increase in the amplitude of excitatory postsynaptic potentials

(EPSPs), was greater than the LTF produced by the standard protocol. Figure parts **c** and **d** are from REF. 85, Nature Publishing Group.

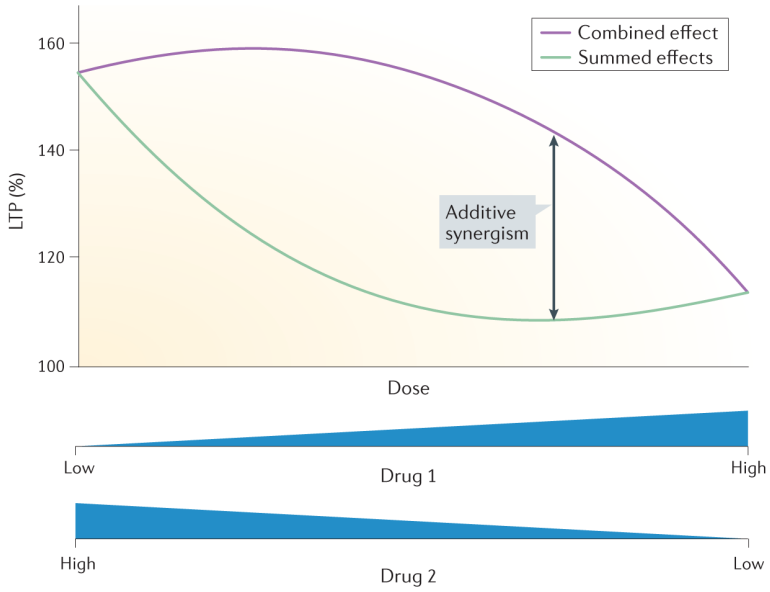

Figure 5. A model predicts that a pair of drugs can act synergistically to enhance LTP

A CREB-binding protein (*Cbp*) mutation impairs hippocampal long-term potentiation (LTP) and impairs learning in mice, and $Cbp^{+/-}$ mice are considered to be a model for aspects of Rubinstein–Taybi syndrome in humans[104]. We developed a model to examine whether drugs could be used to overcome this impairment in LTP. This figure was generated from a series of simulations of the effects of two drugs on the induction of LTP. LTP was modelled as the percentage increase in a synaptic weight variable. In the absence of drugs, simulated LTP induced by a high-frequency tetanic stimulus was strongly impaired. Only a 50% increase in synaptic weight for C$bp^{+/-}$ occurred, compared with an increase in synaptic weight of 148% with non-mutated *Cbp*. The effect of each drug was simply modelled as a change in the value of a kinetic parameter. In this series of simulations, the doses of two drugs — drug 1, a cyclic AMP phosphodiesterase inhibitor, and drug 2, an acetyltransferase activator — were concurrently varied. The effect of drug 1 was simulated by decreasing a rate constant for cAMP degradation, and the effect of drug 2 was simulated by increasing a rate constant for histone acetylation. The ‘dose’ of drug 1 — the amplitude of the rate constant change — was increased, and simultaneously the dose of drug 2 was decreased. Eighty pairs of drug doses were simulated. Both drugs substantially enhanced LTP. For drug 2 alone (left end point of the graph), LTP was 155%, and for drug 1 alone (right end point) LTP was 116%. For both drugs together, with smaller doses of each drug, intermediate LTP amplitudes were observed (combined-effect curve). This series of simulations further shows that additive synergism persists over a substantial range of drug doses. Additive synergism is quantified as the difference (black double arrow) between the LTP simulated when both drugs are applied together (combined effect curve), and the LTP simulated by adding together the effect of the drugs applied individually in separate simulations (summed effects curve).