COGNITIVE LOAD THEORY AND THE USE OF WORKED EXAMPLES AS AN INSTRUCTIONAL STRATEGY IN PHYSICS FOR DISTANCE LEARNERS: A PRELIMINARY STUDY

Kim Guan SAW
School of Distance Education
Universiti Sains Malaysia
Penang, Malaysia

ABSTRACT

This article revisits the cognitive load theory to explore the use of worked examples to teach a selected topic in a higher level undergraduate physics course for distance learners at the School of Distance Education, Universiti Sains Malaysia. With a break of several years from receiving formal education and having only minimum science background, distance learners need an appropriate instructional strategy for courses that require complex conceptualization and mathematical manipulations. As the working memory is limited, distance learners need to acquire domain specific knowledge in stages to lessen cognitive load. This article charts a learning task with a lower cognitive load to teach Fermi-Dirac distribution and demonstrates the use of sequential worked examples. Content taught in stages using worked examples can be presented as a form of didactic conversation to reduce transactional distance. This instructional strategy can be applied to similar challenging topics in other well-structured domains in a distance learning environment.

Keywords: Worked examples, cognitive load theory, instructional strategy.

INTRODUCTION

Open and distance learning provides the working adult with opportunities to upgrade themselves in terms of academic qualifications as well as skills to remain attractive in the labor market. Distance learners are usually adult learners and the transition from a non-student to a student role can be a harrowing experience for some (Rice, 1982). Acquiring domain specific knowledge in stages is an important strategy for these learners to be successful. Higher level undergraduate STEM courses such as physics courses usually require complex conceptualization and mathematical manipulations. Strategies to teach physics courses can be drawn from cognitive load theory (CLT) which views the human cognitive architecture as a natural information processing system (Paas et al., 2003; Sweller, 2012; Wong et al., 2012). Based on a cognitive architecture that consists of a limited working (short-term) memory that interacts with an unlimited long-term memory, CLT requires instructional strategies to engage students in activities that focus on schema acquisition and automation and ensure a processing capacity that learners can handle. Consequently, instruction should be designed in ways to enable working memory to process instructions and construct schema using appropriate levels of cognitive load. Since the learner's working memory load should not be exceeded when processing instruction, instructional strategies should keep the total levels of intrinsic and extrinsic cognitive load to within the working memory limits of the learner. Recent studies showed that there is continuous keen research interest on CLT and instructional strategies appropriate for distance learners (Abeysekera and Dawson, 2015; Asraj et al., 2011; Choi et al., 2014; Chu, 2014; Wong et al., 2012).

Previous studies have shown that in well-structured domains such as physics and mathematics, worked examples can be used as a cognitive load reducing technique (Sweller and Cooper, 1985; VanLehn, 1996). A worked example usually presents a problem based on a principle or theory that has been introduced and provides solution steps that lead to the final answer. Learning from worked examples has been reported to be the preferred mode by novices (LeFevre and Dixon, 1986; Renkl, 2002) and is particularly effective during the initial stages of cognitive skill acquisition (Renkl and Atkinson, 2002; Renkl and Atkinson, 2003). It is particularly helpful to distance learners as it can be considered as a form of guided didactic conversation as proposed by Holmberg (1983) who views effective distance learning as a guided conversation between learner and teacher separated by a transactional distance. Zhu and Simon (1987) found that the entire curriculum of a three-year program in algebra can be taught in two years without performance deficits by using carefully designed sequential worked examples to induce problem-solving mathematical skills even when face-to-face instruction is absent. While worked examples are preferred to problem solving, the extent to which learners benefit from studying the examples is believed to depend greatly on how well they explain the given solutions to themselves (Chi et al., 1989; Salden et al., 2010). Chi et al. (1989) noted that learners who benefited from the study of physics worked examples are those who frequently related the operators to domain principles or principle-based explanations. These learners were also observed to frequently elaborate on the application conditions and goals of operators.

The complex interactions and procedures required in higher level undergraduate physics courses have implications for designing instruction. Teachers as well as instructional designers who use highly complex learning tasks from the start of the course are likely to pile excessive cognitive load on the learners (van Merrienboer et al., 2003). Sweller et al. (1998) rightly posited that working memory is incapable of highly complex interactions unless most of the information with which learners reason has previously been stored in the long-term memory. A good instructional strategy should enable learners to acquire and store domain specific knowledge in stages to lessen cognitive load from complex reasoning processes involving combinations of unfamiliar information at the beginning. A sequence of worked examples that allows the acquisition and storing of domain specific knowledge and an effective construction of schemas is thus suggested. This article revisits the cognitive load theory and illustrates how worked examples can be used as a strategy to teach the topic on the Fermi-Dirac distribution function in Statistical Physics to distance learners at the School of Distance Education (SDE), Universiti Sains Malaysia. The Physics program is one of the eleven undergraduate programs offered to students who possess the *Sijil Tinggi Persekolahan Malaysia* (which is equivalent to the 'A' level qualification) and who wish to pursue their first degree while keeping a full-time job. The students who are mostly in their thirties have left formal education for several years. Being a dual-mode university, Universiti Sains Malaysia maintains a relatively low intake of distance learners to concentrate its resources and to ensure that the degree programs offered via distance learning are equivalent to those offered to conventional face-to-face students. This implies that while the entry requirements are more flexible for students registering for the distance programs, the exit requirements such as the total number of credit hours are similar to the face-to-face students. The flexible entry requirements implies that students with minimum science background may be allowed to register for a science program if they fulfill other conditions.

The SDE has been utilizing web and video conferencing learning environment as the main delivery mechanism for its degree programs. This delivery system allows real time interaction between teacher and student(s), as well as student and student with the capability of presenting teaching materials using software such as Microsoft PowerPoint, simulation packages or video clips. A document camera that is linked to the system is also available should the teacher choose to teach while writing on paper or to demonstrate 3-D objects during the lesson. Recorded web and video conferencing sessions are kept in a learning management system that also stores course contents and activities, assignments as well as assessment grades. In addition, students are given printed materials including

textbooks and in-house publications. Since the beginning of the 2006/07 academic year, STEM courses have been taught in English.

COGNITIVE LOAD THEORY

Cognitive load theory is a useful framework of instructional strategies for complex cognitive domains due to the characteristics and relations between working memory and long-term memory, both of which are structures that constitute human cognitive architecture. It is essentially an instructional theory and is concerned with the learning of complex cognitive tasks where learners are often overwhelmed (Paas et al., 2004). Since the main purpose of instruction is to construct schemas in the long-term memory, teachers and instructional designers should avoid strategies that ignore working memory limitations. For the unaided individual (learner) the working memory is believed to be able to hold about 7 items or chunks of information (Miller, 1956). Information is stored in the long-term memory as schema after being processed by working memory. The information that needs to be learned is known as an element while a schema is a cognitive construct that organizes the elements according to the manner with which they will be dealt with (Sweller, 1994). Schemas can also act as elements themselves in higher order schemas. This implies that acquired schemas can subsequently act as elements when a learner is faced with more complex tasks in the future. Intrinsic cognitive load depends on the content to be learned, and is characterized or determined by the number of interacting elements that the learning task or learning material comprises (Sweller, 1994, 2010). For complex cognitive domains, intrinsic cognitive load may be high and instructional strategy is therefore required for meaningful learning. Larkin et al. (1980) suggest that working memory may contain about 20 elements, given the uncertainty on how information should be packaged into elements. Examples of information stored in the working memory during problem-solving in physics include the current goal (e.g. how to interpret symbols in a particular equation), the status of various quantities, and the assigning of symbols to quantities. While intrinsic cognitive load depends on the inherent complexity of the subject material, the key to reducing intrinsic cognitive load in a difficult subject is to find simpler learning tasks where some interacting elements can be removed (Paas et al., 2003). This is particularly important to distance learners who have left formal education for several years and are beginning to embark on an undergraduate program.

When a learner has gained familiarity with a domain and the need to devote attention to the required cognitive processes is greatly reduced, automation is said to have occurred. Automatic processing occurs when information stored in schemas can be processed automatically without conscious control or effort (Stoica et al., 2011; Sweller, 1994) and thus frees cognitive resources for other activities. Worked examples can be used as an instructional strategy to help learners gain familiarity with a domain and reduce cognitive load.

STATISTICAL PHYSICS AS AN UNDERGRADUATE COURSE FOR DISTANCE LEARNERS

Statistical Physics predicts the average value of the thermodynamic properties based on probabilistic treatment of an atomic model of the system. It is obvious that complete information of a system with a large number of degrees of freedom cannot be obtained and thus statistical consideration can be used to predict the macroscopic properties from the atomic model. Students taking this course are expected to have come from a science background where calculus, probability and permutations have already been taught in high school. In university these students would have studied Newtonian mechanics, quantum mechanics, electrodynamics and thermodynamics in their first and second years before registering for this course. Typical learning outcomes are often stated as the ability to describe the statistical nature of concepts and laws of thermodynamics as well as the ability to use the Boltzmann, Fermi-Dirac and Bose-Einstein distributions to solve problems in some physical systems.

Statistical Physics is a core course for students majoring in Physics at the School of Distance Education. The distance learners from two different batches indicated that part of the difficulty in learning Statistical Physics lies in the high intrinsic load as the subtopics almost always contain abstract concepts and mathematics-related problems where problem-solving skills are required. These subtopics are linked in many ways but unfortunately most sources of information do not give a complete understanding on how they are linked. As distinct elements must be learned separately, the distance learner has to integrate them mentally from different sources. This process inevitably creates a high extraneous cognitive load that is known as the split-attention effect (Mayor and Moreno, 2003). The split-attention effect implies that giving attention to different sources of information increases the extraneous cognitive load on working memory and thus impedes the learning process. One of the topics in Statistical Physics that the distance learners at SDE find hard to understand is the derivation of the FD function. A needs analysis survey involving a total of 40 respondents from two different academic years indicated the following were lacking:

- **an understanding of the abstract concept of the Pauli Exclusion Principle. Specifically, how this principle is translated into the arrangement of particles among energy levels;**
- **a simpler method to derive the FD function without using different sources of techniques and information such as the difficult Lagrange method and the Stirling approximation.**

WORKED EXAMPLES IN THE TEACHING OF PHYSICS

While the worked example effect has been the focus of research in recent years (Cooper, 1990; Leppink et al., 2014; Margulieux and Catrambone, 2016; Mwangi and Sweller, 1998; Stoica et al., 2011; Sweller, 1994), studies on using worked examples as instructional strategies in higher level physics courses are rare. Two previous studies by Cooper (1990) and Chi et al. (1989), involved low level undergraduate physics courses. Cooper (1990) described how a worked example can be used to teach insulation resistance testing and discussed the difference in cognitive load between a split source of information and an integrated format of instruction. The study, however, did not analyze the conceptual structure of the subject content nor identify essential factual and procedural knowledge. Chi et al. (1989) found that the extent to which learners in Mechanics gain from the use of examples depended on how well they explained the solutions to the worked examples to themselves. Knowing how to explain the solutions shown is deemed necessary as worked examples typically contain unclear actions. Learners should therefore compensate for the action gaps in the examples by providing explanations for a particular action that is shown in the worked example.

Worked examples make more efficient use of students' limited cognitive resources than problem-solving (Moreno, 2006). While problem-solving skills in well-structured domains such as higher level physics courses require a large number of schemas as extensive search processes are involved, learning via worked examples is believed to be a practice method that makes more efficient use of students' limited cognitive resources than problem-solving (Moreno, 2006). Worked examples are believed to be more effective as they involve less random processes than problem-solving (Sweller, 2006). They model the process of problem-solving by presenting an example problem and showing the necessary steps towards the final answer to the problem (Renkl, 2002; Renkl and Atkinson, 2002).

The sequential worked-example effect is based on the premise that choosing a simpler learning task that omits some interacting elements can reduce intrinsic cognitive load and therefore facilitate learning. Sequencing the learning material in a simple-to-complex order so that learners avoid the full complexity at the outset is suggested as a way to control intrinsic cognitive load and lessen the instantaneous load (de Jong, 2010; van Merrienboer et al., 2003). For higher level courses, a single worked example for a particular instructional area will probably not result in reducing the cognitive load

(Sweller, 2006). The effectiveness of worked examples is believed to be due to a lengthened 'example phase' where a number of examples are presented before learners are expected to engage in problem-solving (Renkl and Atkinson, 2003). Similarly, Cooper and Sweller (1987) reported that learners who were trained with worked examples or an extended period were better able to solve problems as well as transfer problems compared to those who were subject to conventional problem-solving training.

Sequential worked examples can be utilized to teach the derivation of the Fermi-Dirac (FD) distribution function of a system of particles as this topic can be divided into different stages of presentations. In a system with a relatively small amount of particles and fixed energy the occupation numbers of the energy levels can be calculated easily. The difficulty arises when the number of particles is very large. A general expression for the average occupation numbers is therefore needed. The FD distribution function is an expression that describes the most probable distribution of identical and indistinguishable particles (fermions) that obey the Pauli Exclusion Principle. It gives the number of particles occupying state *i* that has energy E_i. It can be expressed as

$$f_{FD} = \frac{1}{e^{\left(E_i - E_f\right)/kT} + 1}$$

where *k* is the Boltzmann constant, *T* is the absolute temperature, E_i is the energy of the state *i* and E_f is the Fermi energy, which is characteristic of the system being described.

The derivation of the function has a physical significance as it gives the expected number of particles occupying the state *i* that has energy E_i. In other words, the function gives the probability that a state with energy E_i is occupied by a particle. The FD distribution clearly indicates that at the temperature at absolute zero (T = 0 K) particles will fill up available energy states below the characteristic energy level E_f following Pauli Exclusion Principle.

The derivation of the FD distribution function can be also as a sequence of physics principles selected and applied to the problem situation. In a previous study, Sweller (1988) noted that problem solving by means-ends analysis may impede learning as it demands a high cognitive processing capacity that is consequently inaccessible for schema acquisition especially for novices. The study shows that expert and novice problem solvers in physics use different strategies, with novices using the means-ends analysis. Novices were reported to work backwards from the goal where subgoals were set. The backward process were performed until they obtain the required equations for the subgoals were obtained. With this achieved, the backward process were consequently reversed and a forward-working sequence commenced. As the domain specific knowledge and the memory of problem states of the experts were different, they could proceed with the forward-working sequence with the appropriate choice of equations that leads to the goal. Since distance learners with minimum science background are considered novices, worked examples may be a better instructional strategy than problem solving.

WORKED EXAMPLES AS AN INSTRUCTIONAL STRATEGY TO TEACH THE FD DISTRIBUTION FUNCTION

Learners can be instructed on the derivation of FD distribution function via different learning tasks. Most of these tasks require highly complex mathematical skills such as calculus, the use of Stirling's approximation as well as the method of Lagrange multipliers. A usual route to teach FD distribution function starts at imagining a three-dimensional virtual space such that a state of the particle is represented by a point in this space. All possible states of the particle are located only in 1/8 of the total space for positive values on the three orthogonal axes. Some of the major steps of determining the density of states are listed as follows:

a) **Relate the kinetic energy of a particle to its momentum**
b) **From (a) the expressions** $p^2 = 2mE$ **and** $dp = \sqrt{2}m^{3/2}E^{1/2}dE$ **are obtained**

c) **Obtain the density of states which is expressed as** $g(E)dE = \frac{4\pi V}{h^3}\sqrt{2}m^{3/2}E^{1/2}dE$

d) **Modify the density of states for the spin orientation to be taken into consideration to obtain the actual density of states which is** $dS = \frac{8\pi V}{h^3}\sqrt{2}m^{3/2}E^{1/2}dE$

Further steps involving the partition function, the Helmholtz free energy and the method of Lagrange multipliers are required before the learner can finally obtain the FD distribution function. These learning tasks therefore involve a high cognitive load and element interactivity. Using highly complex learning tasks from the beginning will affect learning, performance and motivation negatively as it imposes excessive cognitive load on the learners (van Merrienboer et al, 2003; Sweller et al., 1998).

A better learning task should present a lower intrinsic cognitive load and starts with understanding the difference between permutations and combinations of objects before moving on to arranging indistinguishable particles among the energy levels of a system with a fixed total number of particles and energy. This route is particularly helpful for distance learners who have minimum science background. The FD statistics can be derived from simple concepts associated with the arrangements of indistinguishable particles subject to the Pauli Exclusion Principle. The procedures to be applied are basically the combination of particles in the different energy states in the various available energy levels subject to Pauli Exclusion Principle. For the derivation of the FD distribution function, the learning task takes the simpler route of comparing two similar systems of particles subject to Pauli Exclusion Principle in which the second system has one less particle than the first.

This learning task therefore lets learners start their work on relatively simple content and progress towards more complex content. It also omits many interacting elements that are associated with the earlier learning tasks that rely mainly on the density of states and the method of Lagrange multipliers. In view of this, teachers as well as instructional designers can use sequential worked examples as a strategy to reduce cognitive load and facilitate learning in higher level undergraduate physics courses. The worked examples should be able to demonstrate how to apply the required procedures to the specific steps leading to the derivation of the FD distribution function.

In this learning task it is clear that a sequence of worked examples should consist of:

i. **example(s) that differentiate permutation and combination;**
ii. **example(s) showing the distribution of indistinguishable particles in systems with a fixed amount of energy and a small number of particles for non-degenerate energy levels;**
iii. **example(s) showing the distribution of particles obeying the Pauli Exclusion Principle for a system with the degeneracy of energy levels;**
iv. **example(s) showing the derivation of the FD statistics;**
v. **example(s) showing the derivation of the FD distribution function based on the comparison of two systems of particles where the second system has one less particle at an arbitrary energy level.**

Although the worked examples are sequenced, they should all be integrated examples rather than split-source examples. This is important to note that Mwangi and Sweller (1998) found the latter examples difficult to comprehend.

Worked Example on Permutations and Combinations

Question

Five boxes are labeled A, B, C, D and E. Find the number of ways the three boxes can be chosen if

i. **the order is important**
ii. **the order is irrelevant**

Solution

This is a question that involves the concepts of permutation and combination. In the case of permutation, the order should be taken into consideration but in combination the order is irrelevant.
If order is important then ABC is considered a different arrangement from ACB and so on. For permutation, we need to use the formula

$$^{n}P_{r} = n(n-1)(n-2)\ldots(n-r+1)$$

where *n* is the total number of available boxes and *r* is the number of boxes that will be chosen.

The answer for (i) is

$$^{5}P_{3} = 5\times4\times3 = 60$$

When the order of choice is not relevant: ABC, ACB, etc are the same selection. For combination, the formula to use is

$$^{n}C_{r} = \frac{n!}{r!\times(n-r)!}$$

The answer for (ii) is therefore

$$^{5}C_{3} = \frac{5\times4\times3}{3!} = 10$$

Worked Example on the Distribution of Indistinguishable Particles for Non-degenerate Energy Levels

Question

A system consists of three indistinguishable particles and a fixed total energy of 3ϵ. Show the possible macrostates i.e. show how the particles can be placed in any of the four non-degenerate energy levels: 0, ϵ, 2ϵ and 3ϵ.

Solution

The macrostate is defined by the energy. This will then allow us to have the following possibilities, each defining a microstate of the system:

i. **one particle in the energy level state of 3ϵ and the remaining two in the energy level state 0;**
ii. **one particle in the energy level state of 2ϵ, one in the energy level state of ϵ while the last particle is in the energy level state of 0.**
iii. **all three particles are in the energy level state of ϵ.**

	I	II	III
3ϵ	•		
2ϵ		•	
ϵ		•	•••
0	••	•	

Figure 1.The possible macrostates of a system with three indistinguishable particles with a total energy of 3ϵ.

In this case we have three ways to 'arrange' the particles among the energy levels as shown in Figure 1. A macrostate is defined by the specification of the number of particles in each energy level. Thus in this isolated system consisting of three particles with a total energy of 3ϵ, there are three macrostates. Since the energy levels do not degenerate, there is only one energy state in each energy level. Moreover, since the particles are indistinguishable (i.e. the particles are not labeled), interchanging the particles will not result in a different distribution.

Worked Example Comparing the Differences on the Distribution When Particles in a System with Degenerate Energy Levels are Subject to Pauli Exclusion Principle.

Question

A system consists of three indistinguishable particles and a fixed energy of 3ϵ. The degeneracy of the energy level is 2. Show the possible macrostates if (a) the particles are not subject to Pauli Exclusion Principle (b) the particles are subject to Pauli Exclusion Principle.

Solution

Since the degeneracy of the energy levels is 2, there are two energy states in each energy level. Moreover, since the particles are indistinguishable (i.e. the particles are not labeled), interchanging the particles among the states or energy levels will not result in a different distribution.

For (a), the particles are not subject to Pauli Exclusion Principle and thus more than one particle can be placed in an energy state. There are three possible macrostates (labeled as macrostates I, II and III). In macrostate I, one particle is placed in energy level 3ϵ and two in the energy level at ground state (ϵ=0) to meet the condition of a total number of three particles and a total energy of 3ϵ. There are altogether 6 possible microstates among the energy states as shown in Figure 2. Microstates 1 – 4 exist because two particles can exist at the same time in one energy state as they are not subject to the Pauli Exclusion Principle.

	I	1	2	3	4	5	6
3ϵ	•	•		•		•	
			•		•		•
2ϵ							
ϵ							
0	••	••	••			•	•
				••	••	•	•

Figure 2.The distribution of particles not subject to Pauli Exclusion Principle for macrostate I.

In macrostate II, there is one particle each in energy levels 0, ϵ and 2ϵ, giving the total energy 3ϵ as required. There are eight microstates in this macrostate as shown in Figure 3.

	II	1	2	3	4	5	6	7	8
3ϵ									
2ϵ	•	•	•			•	•		
				•	•			•	•
ϵ	•	•		•		•		•	
			•		•		•		•
0	•					•	•	•	•
		•	•	•	•				

Figure 3.The distribution of particles not subject to Pauli Exclusion Principle for macrostate II. (Note that the distribution for particles that obey the Pauli Exclusion Principle is similar).

Next, we look at macrostate III. In macrostate III, there are three particles in the energy level ϵ, which gives a total energy of 3ϵ. In this macrostate there are four microstates as shown in Figure 4.

	III	1	2	3	4
3ϵ					
2ϵ					
ϵ	•••	•••	••	•	
			•	••	•••
0					

Figure 4.The distribution of particles not subject to Pauli Exclusion Principle for macrostate III.

For (b) the number of macrostates is reduced since the particles are now subject to the Pauli Exclusion Principle. Only macrostates I and II are possible. However, unlike the previous example for macrostate I, microstates 1 – 4 in Figure 3 are now not possible as there are two particles in one energy state. Thus only microstates 5 and 6 in Figure 3 are allowed. Figure 5 shows the distribution of particles subject to Pauli Exclusion Principle for macrostate I. Macrostate III is not allowed as there are three particles in energy level 3ϵ that has only two energy states.

	I	5	6
3ϵ	•	•	
			•
2ϵ			
ϵ			
0	••	•	•
		•	•

Figure 5.The distribution of particles subject to Pauli Exclusion Principle for macrostate I.

Worked Example on Obtaining the FD Statistics

Question

Obtain the FD statistics that describes the system of particles which obey the Pauli Exclusion Principle.

Solution

A system of particles obeying the Pauli Exclusion Principle is described by a set of occupation numbers that specifies the number of particles in the energy levels. We note that in this case, one-particle states are only permitted to have occupation numbers of 0 or 1. Each energy level contains a number of energy states, g_j. Let us suppose in an arbitrary level j, a certain arrangement of n_j particles is such that in State 1, there is only one particle, p. In this case, alphabets are temporarily attached to the particles although they are indistinguishable. Not more than one particle may occupy each permitted energy state, which implies some energy states may be empty. A possible arrangement for the particles where only the first few states are represented is shown in Figure 6.

State 1	(p)
State 2	
State 3	(q)
State 4	(r)
State 5	

Figure 6. A possible arrangement of particles subject to Pauli Exclusion Principle for an arbitrary energy level with several energy states.

The first particle p is "spoilt for choice" as there are g_j possible locations. The second particle q has only (g_j – 1) possible locations and the third particle r has (g_j – 2) locations, and finally the last particle has very little choice i.e. (g_j–n_j + 1) locations. The total number of ways in which a particular sequence of n_j alphabets can be allocated to the available g_j states is simply

$$\frac{g_j!}{\left(g_j - n_j\right)!}$$

Next, we remove the alphabets and this number of arrangements is divided by the number of arrangements of n_j particles among themselves i.e.

$$^{g_j}C_{n_j} = \frac{g_j!}{n_j!(g_j - n_j)!}$$

Finally, we should take into consideration all energy levels of the total number of indistinguishable particles, the total number of arrangements, *t*, becomes

$$t = \prod_j \frac{g_j!}{n_j!\left(g_j - n_j\right)!}$$

This is the FD statistics. Therefore for a particular energy level E_j in a system of fermions (i.e. particles with half-integer spin) there is a definite number of g_j states that will possess this energy. Due to the Pauli Exclusion Principle, the maximum number of fermions that can occupy this energy level will thus be g_j.

Worked Example on the Derivation of FD Distribution Function

Question

Derive the FD distribution function.

Solution

A distribution function is a general expression for the average occupation numbers when the total number of particles *is very large*. The goal is to find the average number of particles (fermions) in a state with a certain amount of energy as a function of temperature. The FD distribution function can be derived by comparing two systems (assemblies) of particles that obey the Pauli Exclusion Principle in which the second system has one less particle than the first. The two systems have the same energy levels with the same degeneracy. This method of comparison will help us avoid using Lagrange multipliers.

Suppose the first system consists of five indistinguishable particles and a total energy of 4ϵ. The degeneracy of the energy level is 3. This implies that the second system should only consist of four particles. Three possible macrostates of the first system are shown in Figure 7. The degeneracy level is represented by the 'trays'.

Macrostate, k / Energy Level, ϵ_j	1	2	3
$\epsilon_3 = 3\epsilon$			•
$\epsilon_2 = 2\epsilon$	•	••	
$\epsilon_1 = \epsilon$	••		•
$\epsilon_0 = 0$	••	•••	•••

Figure 7. The distribution of particles for the first system.

From the FD statistics, the statistical weight of the first macrostate is

$$W_{k=1} = \frac{3!}{2!(3-2)!} \times \frac{3!}{2!(3-2)!} \times \frac{3!}{1!(3-1)!} = 27$$

Here *k*=1 denotes the first macrostate while the degeneracy of the energy levels, *g_j*, is 3. The number of particles in each energy level, *n_j*, can be seen in Fig. 7. Similarly, the statistical weights for the second and third macrostates are

$$W_{k=2} = \frac{3!}{3!(3-3)!} \times \frac{3!}{2!(3-2)!} = 3$$

$$W_{k=1} = \frac{3!}{3!(3-3)!} \times \frac{3!}{1!(3-1)!} \times \frac{3!}{1!(3-1)!} = 9$$

Note that 0! = 1.

The total number of microstates is

$$\Omega = \sum_k W_k = 27 + 3 + 9 = 39$$

The average occupation number for a certain energy level can also be obtained from this simple system. For instance, the average occupation number in level 1 is

$$\langle N_1 \rangle = \frac{1}{\Omega} \sum_k N_{1k} W_k = \frac{(2\times 27) + (1\times 9)}{39} = 1.615$$

An arbitrary level *l* is chosen where one particle is taken out for the second system. If energy level 2 is chosen then the second system will have a total energy of 2ϵ, and in every macrostate of the first system the occupation number of all energy levels with the exception of level l = 2 will be the same i.e. $N_j^* = N_j$ where j≠l. (If $j = l$ then for a particular macrostate k, $N_{lk}^* = N_{lk} - 1$).

Only two macrostates are possible as shown in Figure 8.

Macrostate, k / Energy Level, $\in_j$	1		2
$\in_3 = 3\in$			
$\in_2 = 2\in$			•
$\in_1 = \in$	••		
$\in_0 = 0$	••		•••

Figure 8. The distribution of particles for the second system.

We denote N_{jk} and N_{jk}^* as the occupation numbers of level *j* in macrostate *k* in the first and second systems, respectively. There are no macrostates in the second system that corresponds to the states in the first system where the level *l* is not occupied. For such macrostates, the statistical weights represented by W_{lk}^* are therefore nil.

Since $N_{lk}! = N_{lk}\left(N_{lk}-1\right)! = N_{lk}N_{lk}^*!$ the ratio of the statistical weights of corresponding macrostates in the first and second systems is

$$\frac{W_{lk}^*}{W_{lk}} = \prod_j \frac{\left(g_j - N_{jk}\right)!\left(N_{jk}\right)\left(N_{jk}-1\right)!}{\left(g_j - N_{jk}^*\right)!N_{jk}^*!} = \prod_j \frac{\left(g_j - N_{jk}^* - 1\right)!N_{jk}}{\left(g_j - N_{jk}^*\right)\left(g_j - N_{jk}^* - 1\right)!} = \frac{N_{lk}}{\left(g_l - N_{lk}^*\right)}$$

(We are not really interested in this ratio of $\frac{W_{lk}^*}{W_{lk}}$. To obtain the FD distribution function we are actually interested in getting the ratio of the average number of particles in a state at energy *E*$_j$ to the number of states with this energy, which is $\frac{\langle N_j \rangle}{g_j}$. But we have to go on with several more steps before we can obtain the answer that we want).

For systems with a very large number of particles, the removal of a particle from an energy level will not make any significant change in the average occupation number of that level. Thus the average occupation number of the arbitrary energy level of the first system can be assumed to be equal to that of the second system. The summation over all macrostates leads to

$$g_l\Omega_l^* - \langle N_l^* \rangle \Omega_l^* = \langle N_l \rangle \Omega_l$$

where $\Omega = \sum_k W_k$ and $\Omega^* = \sum_k W_k^*$ are the total number of microstates in the first and second systems, respectively.

We now need to take the natural logarithms of both sides and together with the definition of entropy we have

$$\ln \frac{\langle N_l \rangle}{g_l - \langle N_l^* \rangle} = \ln \Omega_l^* - \ln \Omega_l = \frac{\Delta S}{k}$$

Remember that for the microcanonical ensemble, *S = k* ln Ω.

Based on Gibb's free energy, the difference in entropy which is related to the temperature and the difference in energy can be expressed as

$$T\Delta S = \Delta U - u\Delta N$$

where *μ* is the chemical potential for each particle. Since there is only one particle less in the second system used for comparison, $\Delta N = -1$ **and** $\Delta U = -\varepsilon_l$**.**

The difference in entropy can be written as

$$\Delta S = \frac{\mu - \varepsilon_l}{T}$$

For systems with a large number of particles and since level *l* was selected arbitrarily and can be representing any energy level *j*, the relation between the occupation number, entropy and degeneracy, chemical potential μ and temperature *T* can then be expressed as

$$\ln \frac{\langle N_j \rangle}{g_j - \langle N_j \rangle} = \frac{\mu - \varepsilon_j}{kT}$$

(The chemical potential μ is sometimes referred to as the Fermi level). Removing the natural logarithm from the equation leads to

$$\frac{\langle N_j \rangle}{g_j - \langle N_j \rangle} = e^{\frac{\mu - \varepsilon_j}{kT}}$$

Taking the reciprocal

$$\frac{g_j - \langle N_j \rangle}{\langle N_j \rangle} = \left(e^{\frac{\mu - \varepsilon_j}{kT}} \right)^{-1} = e^{\frac{\varepsilon_j - \mu}{kT}}$$

$$\frac{g_j}{\langle N_j \rangle} = e^{\frac{\varepsilon_j - \mu}{kT}} + 1$$

Finally,

$$\frac{\langle N_j \rangle}{g_j} = \left[e^{\frac{\varepsilon_j - \mu}{kT}} + 1 \right]^{-1}$$

We have obtained the ratio that we wanted. This is the FD distribution function which is the average number of particles (e.g. fermions) in a state. However, we do not have to denote the ratio as it is well-understood. Thus the FD distribution function can be expressed as

$$f_{FD} = \frac{1}{e^{(E_i - E_f)/kT} + 1}$$

where *k* is the Boltzmann constant, *T* is the absolute temperature, E_i is the energy of the state *i* and E_f is the Fermi level.

DISCUSSIONS

The worked examples suggested here can help the distance learners to focus on how abstract principles are used to solve problems that are related to the derivation of the FD distribution function. In the first worked example, there is an application of an elementary probability principle. While this worked example appears to be at a fundamental level of understanding particle arrangement it is particularly useful to distance learners with minimum science background. They need to know the principle which includes the permutation (and combination) principle as well as the meaning of *n* and *r* besides the types of problems to which the formula applies. Rather than having concrete features that inhibit transfer to new problems, the worked examples presented here as a sequence promote conceptual understanding. Distance learners who have been exposed to this sequence of worked examples should also be able to derive the Bose-Einstein and the Maxwell-Boltzmann distribution functions. The teaching of the FD distribution function is not limited to these worked examples. While these sequential worked examples are prepared according to the findings of the needs analysis survey their potential use is obviously not limited to students in the distance learning mode. Teachers can also modify these worked examples to expand the scope of the subtopics required. However, in recent years, there has been some concern that when learners rely too heavily on worked examples, they are unable to glean conceptual information about the procedure from these concrete examples and thus are unable to transfer the necessary skills to new problems (Catrambone, 1998; Margulieux and Catrambone, 2016; Renkl, 2002). The use of sequential worked examples is believed to be able to overcome over-reliance on demonstrated solutions as learners need to relate the operators to domain principles as well as link all the examples in their attempt to understand the main content that is being taught. Sequential examples can be considered as a form of scaffolding practice within the cognitive load theory (Rosenshine and Meister, 1992; van Merrienboer et al., 2003) to assist learners bridge the gap between their current abilities and the intended goal. As scaffolds, the worked examples that are presented as a sequence provide temporary support and allow the distance learners to participate at an ever increasing level of competence as they move from a lower level worked example to a higher level one. In particular, the distance learners move from simpler tasks of permutations and combinations to the distribution of indistinguishable particles in energy levels in systems that obey the Pauli Exclusion Principle as well as those that do not. In the last worked example, the distance learners are exposed to a higher level task that requires complex mathematical manipulations to derive the FD distribution function. Avoiding a single long example with action gaps should help the distance learners to generate explanations for the actions shown and subsequently relate these to the scientific principles involved.

CONCLUSION

This article suggests that choosing a simpler learning task that minimizes interacting elements and the use sequential worked examples can be used as an instructional strategy in physics for distance learners. Well-arranged sequential worked examples can help distance learners avoid the split-attention effect as they do not have to give attention to sources of information from various sources. A simpler learning task and well-arranged sequential worked examples are expected to lower the intrinsic and extraneous cognitive load. The strategies and examples presented here can be tested in an empirical study by teachers and researchers in similar disciplines in a distance learning environment.

ACKNOWLEDGEMENT: **This work is partially supported by the Research University Grant (1001/PJJAUH/310456). The author would like to thank Universiti Sains Malaysia for this grant. The grant does not lead to any conflict of interest.**

BIODATA and CONTACT ADDRESSES of AUTHOR

Kim Guan SAW is an Associate Professor at the School of Distance Education, Universiti Sains Malaysia. He obtained his PhD in Physics from the University of Melbourne, Australia, in 2003. He also holds a Bachelor of Science (Physics) with Education (Hons) degree and teaches undergraduate physics courses to distance learners in Malaysia. His research interests include instructional design for distance learning, interaction in virtual learning environment, physics education as well as employability of distance learners. He has published more than 50 articles in international peer-reviewed journals, conference papers as well as research monographs in the fields of both physics and open distance education. He is also a reviewer for international and national journals in physics and open distance education. He has supervised postgraduate students and examined Master's and PhD theses in both these fields.

Kim Guan SAW
School of Distance Education
Universiti Sains Malaysia, 11800 Penang, Malaysia
Phone: +604 6534564,
E-mail: kgsaw@usm.my

REFERENCES

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development, 34(1)*, 1 – 14.

Asraj, A.A., Freeman, M., & Chandler, P.A. (2011). Considering cognitive load theory within e-Learning environments. *PACIS 2011 Proceedings Queensland University of Technology*, 1 – 13.

Catrambone, R. (1998). The subgoal learning model: creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, *127*(4), 355 - 376.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cognitive Science*, *13*(1), 145 – 182.

Choi, H.H., van Merrienboer, J.J.G., & Paas, F. (2014). Effects of the physical environment on cognitive load learning: Towards a new model of cognitive load. *Educational Psychology Review. 26*, 225 – 244.

Chu, H.C. (2014). Potential negative effects of mobile learning on students' learning achievement and cognitive load – A format assessment perspective. *Educational Technology & Society, 17(1),* 332 – 344.

Cooper, G. (1990). Cognitive load theory as an aid for instructional design. *The Australian Journal of Educational Technology*, *6*(2), 108-113.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*(4), 347 – 362.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, 38(2), 105-134.

Holmberg, B. (1983). Guided didactic conversation in distance education. In D. Sewart, D. Keegan, and B. Holmberg (Eds.), *Distance Education: International Perspectives* (p. 115). London: Croom Helm.

Larkin, J., McDermott, J., Simon, D., & Simon, H. (1980). Models of competence in solving physics problems. *Cognitive Science*, *4*(4), 317-348.

LeFevre, J.A., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction*, *3*(1), 1-30.

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C.P.M., van Merrienboer, J.J.G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32-42.

Margulieux, L.E., & Catrambone, R. (2016). Improving problem solving with subgoal labels in expository text and worked examples. *Learning and Instruction*, *42*, 58-71.

Mayer, R.E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90(2)*, 312-320.

Miller, G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Moreno, R. (2006). When worked examples don't work: Is cognitive load theory at an Impasse? *Learning and Instruction*, *16*(2), 170-181.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction*, *16*(2), 173- 99.

Paas, F.G.W.C, & van Merrienboer, J.J.G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *Journal Educational Psychology*, *86*(1), 122-133.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1-4.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instruction Science*, *32*(2), 1-8.

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction*, *12*(5), 529-556

Renkl, A., & Atkinson, R.K. (2002). Learning from examples: Fostering self-explanations in computer-based learning environments. *InteractiveLearning Environment*, *10*(2), 105-119.

Renkl, A., & Atkinson, R.K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, *38*(1), 15-22.

Rice, J. (1982). Spouse support: Couples in educational transition. *Lifelong Learning: The Adult Years*, 6(1), 4 – 6.

Rosenshine, B.V., & Meister, C. (1992). The use of scaffolds for teaching less structured cognitive tasks. *Educational Leadership*, *49*(7), 26-33.

Salden, R.J.C.M., Koedinger, K.R., Renkl, A., Aleven, V., & McLaren, B.M. (2010). Accounting for beneficial effects of worked examples in tutored problem solving. *Educational Psychology Review*, *22*, 379-392.

Stoica, D., Paragina, F., Paragina, S., Miron, C.,& Jipa, A. (2011). The interactive whiteboard and the instructional design in teaching physics. *Procedia Social and Behavioral Sciences*, *15*, 3316-3321.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312.

Sweller, J. (2006). The worked example and human cognition. *Learning and Instruction*, *16*(2), 165-169.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review*, *22*(2), 123-128.

Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. Harris, S. Graham, & T.Urdan (Eds.) *APA Educational Psychology Handbook, Vol. 1* (pp. 295-325). Washington, D.C.: American Psychological Association.

Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59-89.

Sweller, J., van Merrienboer, J.J.G., & Paas. F,G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, *47*,513-539.

van Merrienboer, J.J.G., Kirschner, P.A.,& Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5-13.

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449-457.

Zhu, X, & Simon, H. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, *4*(3), 137-166.