According to studies done in 2013, the proportion of U.S. adults who are functionally illiterate, or unable to read or write in any practical capacity, is as high as 14% (Literacy Project Foundation). This figure includes 24% of the adult U.S. black population, 41% of the Hispanic population, and a staggering 70% of the prison population (U.S. Department of Education, 2015). Students struggle as well, with up to 64% of eighth graders unable to read at their reading level (National Center for Education Statistics, 2013). Naturally, the difficulties faced by this low-literacy population spread across a range of circumstances, making it difficult to find employment, receive medical care, obtain an education, or navigate many other aspects of daily life. With this in mind, the long-term objective of this research proposal is to create tools to improve education in adult literacy and in other academic fields. Specifically, this proposal submits two such tools: an educational system called reMind and an add-on sentence-comprehension module called CAPITAL Passages.

Primarily, reMind will be a system that allows professors to deliver practice and testing questions to students, who will have the ability to practice study materials at their own pace. An external team will implement the bulk of this system, whose key algorithmic distinction will be the spaced repetition of study materials with a deadline. Meanwhile, the portion of this project implemented by this team will involve the development of two user interfaces. One will be a web interface for students and instructors to collaborate and amass a collection of questions over many academic disciplines, and the other will be a mobile interface for students to practice materials and receive immediate feedback about the quality of their responses.

More central to the goal of improving literacy is the CAPITAL Passages system, a question generation module focused on reading comprehension. The ultimate goal of Passages will be to accept a passage of English text and use it to generate coherent comprehension questions from the sentence to entire-passage semantic level. Since Passages will feed questions into reMind, its question-generation algorithms will also need to produce the correct answers to their questions (and incorrect distractor answers when appropriate) in order for reMind to evaluate student responses. Lastly, Passages will also ideally incorporate a classifier that can evaluate whether a generated question makes sense, and later whether instructors actually choose certain questions in practice.

The **intellectual merit** of the project is the challenge of creating the linguistic algorithms for Passages, which will further the field of natural language processing by expanding on current question generation abilities. Passages will integrate multiple sentences into a single question and retain the correct answers to such questions even when question creation becomes complex. Passages will combine different types of question generation—such as true/false and multiple choice questions—into the same tool in order to generate more and higher-quality questions from the same passage. Finally, incorporating machine learning research will enable the system to develop a rigorous classifier that can distinguish questions that make sense (grammatically and otherwise) from those that do not. With these features, Passages will test reading comprehension in a deeper and more rigorous way than a system that simply asks factual questions about one sentence. Potential benefits of this project will be in the fields of natural language processing and education.

The **broader impact** of the project lies in its potential to improve adult literacy and existing education schemes. Through automatic generation, the project will provide a low-maintenance system which can test and train whole-passage comprehension. The project will also allow low-literacy individuals to build literacy-related life skills, which are often a barrier to pursuing education. These improved skills will lead to increased ability of these groups (often impoverished, ethnic minority, or immigrant populations) to participate more fully in education and industry. Furthermore, the reMind system will be able to amass, generate (through modules such as Passages), and store databases of academic questions to be shared among instructors. While the development of these systems will be specific to the English language, some attention to linguistic differences would allow the algorithms to be adapted for other languages.